VOLUME 12, 1961

732

BRITISH JOURNAL OF

Applied Physics

EDITOR: **A. C. STICKLAND,** M.SC., PH.D.

Editor and Deputy Secretary of
The Institute of Physics and The Physical Society

THE INSTITUTE OF PHYSICS AND
THE PHYSICAL SOCIETY
LONDON

THE INSTITUTE OF PHYSICS AND THE PHYSICAL SOCIETY

Officers and Members of the Council, 1961–62

President	Sir John Cockcroft, O.M., K.C.B., C.B.E., M.A., D.Sc., F.Inst.P., F.R.S.
Vice-Presidents	V. E. Cosslett, M.A., M.Sc., Ph.D. Professor R. W. Ditchburn, M.A., B.Sc., Ph.D., F.Inst.P. W. H. Taylor, M.A., D.Sc., F.Inst.P. J. Topping, M.Sc., Ph.D., F.Inst.P.
Honorary Secretary	C. G. Wynne, B.A., Ph.D., F.Inst.P.
Honorary Treasurer	J. Taylor, M.B.E., D.Sc., F.R.I.C., F.Inst.P.
Ordinary Members of Council	Professor C. C. Butler, B.Sc., Ph.D., A.Inst.P. E. R. Davies, O.B.E., B.Sc., F.Inst.P. Professor B. H. Flowers, M.A., D.Sc. J. M. A. Lenihan, M.Sc., Ph.D., A.M.I.E.E., F.Inst.P. G. R. Noakes, M.A., F.Inst.P. Professor C. F. Powell, M.A., Sc.D., F.R.S. F. A. Vick, O.B.E., Ph.D., A.M.I.E.E., F.Inst.P. Professor D. A. Wright, D.Sc., F.R.A.S., F.Inst.P.
Representing Branches	M. R. Hopkins, M.Sc., Ph.D., F.Inst.P. Professor A. M. Taylor, M.A., Ph.D.

Executive Officers

Secretary	H. R. Lang, B.Sc., Ph.D., A.R.C.S., F.Inst.P.
Editor and Deputy Secretary	Miss A. C. Stickland, M.Sc., Ph.D.
Deputy Secretary	N. Clarke, B.Sc., F.Inst.P.

British Journal of Applied Physics

Editorial Board

J. M. A. Lenihan, M.Sc., Ph.D., A.M.I.E.E., F.Inst.P. (*Chairman*)
Professor R. W. Ditchburn, M.A., B.Sc., Ph.D., F.Inst.P. (*Ex-officio member*)
W. Hirst, B.Sc., Ph.D., F.Inst.P.
C. A. Hogarth, B.Sc., Ph.D., F.Inst.P.
J. W. Menter, M.A., Ph.D., F.Inst.P.
A. C. G. Menzies, M.A., D.Sc., F.Inst.P.
Professor D. A. Wright, D.Sc., F.R.A.S., F.Inst.P.

CONTENTS OF VOLUME 12

PAGE

JANUARY 1961

EDITORIAL 1

CONFERENCE REPORT

Summarized Proceedings of a Conference on Stress Analysis—University College, of North Staffordshire, April 1960. By C. D. POMEROY 3

PAPERS

Improvements of photoelastic technique for strain measurement on flat surfaces. By G. CLYNE, H. FESSLER and R. W. WILSON 8

owth of crystals in glass. By J. G. MORLEY . . 10

thod for calculating boundary value problems in heat conduction for the cylindrical cavity and the half space, by means of convolution integrals. By D. W. JORDAN 14

erferometric measurement of small angular displacements. Part 2. The double-passed Jamin interferometer. By P. HARIHARAN and D. SEN 20

erpretation of ultrasonic echo amplitude. By J. C. V. RUMSEY 25

RESPONDENCE

asurement of the tensile strength of brittle materials. From A. ORMEROD, R. BERENBAUM and I. BRODIE . 30

ES AND NEWS

and comments 31

FEBRUARY 1961

IAL ARTICLE

sign of scales for industrial instruments. By A. J. MADDOCK 33

RS

hing and polishing studies on magnesium oxide single rystals. By T. K. GHOSH and F. J. P. CLARKE . . 44

se method for measuring Kerr's constant. By A. IEKARA and R. KONOPKA 50

BOOKS SECTION

iews and notices 53–64

RS (*continued*)

estigation of periodic discontinuities in the mutual otential energy of molecules of water and some polyydric alcohols. By M. M. QURASHI and A. K. M. HSANULLAH 65

ft position indicator for use with journal bearings. By J. HUNTER and C. J. HUGHES 73

CORRESPONDENCE

Dynamic measurement of elasticity using resonance methods. From I. EVANS and G. BRADFIELD . . 80

The influence of the method of demagnetization on the permeability in soft magnetic materials. From A. SMOLINSKI 81

NOTES AND NEWS

Notes and comments 83

MARCH 1961

SPECIAL ARTICLE

Galvanomagnetic effects and their application. By C. HILSUM 85

PAGE

PAPERS

The preparation of thin films of germanium and silicon. By B. A. IRVING 92

Capillary viscometers of negligible kinetic energy effect. By W. A. CAW and R. G. WYLIE 94

Internal friction in fibre assemblies. By J. D. HUFFINGTON 99

Narrow-beam echo-ranger for fishery and geological investigations. By M. J. TUCKER and A. R. STUBBS . 103

Spectral emissivity of 99·7% aluminium between 200 and 540° C. By P. M. REYNOLDS 111

Heating of metallic foils in an electron microscope. By B. GALE and K. F. HALE 115

The frictional properties of lignum vitae. By K. G. MCLAREN and D. TABOR 118

Barium–carbon monoxide system: an hypothesis of Arizumi and Kotani. By T. A. G. GIORGI and S. ORIGLIO 120

NOTES AND NEWS

Notes and comments 123

APRIL 1961

CONFERENCE REPORT AND ASSOCIATED PAPERS

The Joint Conference of the Non-Destructive Testing Group and the Société Française de Métallurgie. By K. W. ANDREWS 125

The structure of non-destructive testing. By K. W. ANDREWS 127

The investigation of surface deformation. By L. GRUNBERG, D. SCOTT and K. H. R. WRIGHT . . 134

Recent advances in magnetic analysis. By K. HOSELITZ 141

Frequency dependence of induced currents (abstract). By P. GRANEAU 146

PAPERS

Rubber models of yarns and cords; the 'doubling' of single rods. By N. WILSON and L. R. G. TRELOAR . . 147

Demagnetization of ferromagnetic particles. By F. SMITH 155

Fields external to open-structure magnetic devices represented by ellipsoid or spheroid. By H. CHANG . 160

Significance of the properties of materials in the packing of real spherical particles. By J. C. MACRAE and W. A. GRAY 164

Porosity and surface area of a granular bed from measurements of the flow of air through the bed and measurements of the granular shape factors. By W. M. JONES and D. MILLS 172

Measurement of thickness of thin transparent films using fluorescence. By B. P. BUNT 175

Conformal transformations by analogue computer. By E. J. JOSS and D. S. ROSS 178

Method for the consistent transformation of physical equations into non-dimensionally parametric form. By J. C. AMSON 180

The determination of stress concentrations with an electrolytic tank model. By T. W. G. CALVERT . . . 184

The coupling factor of piezoelectric ceramic disks. By D. S. CAMPBELL and A. M. MACSWAN . . . 188

The copper staining of p–n–p alloyed junction transistor sections. By S. RATCLIFFE and J. E. HUGHES . . 193

CORRESPONDENCE

Comments on the A.S.T.M. Powder Data File—6. From J. W. HUGHES, ISABEL E. LEWIS and A. J. C. WILSON . 194

DETACH AFTER SEWING ONLY

PAGE

NOTES AND NEWS

Notes and comments 195

MAY 1961

INVITED ARTICLES

Amplification through stimulated emission—the MASER. By R. A. SMITH 197
The physics of computer elements. By C. N. W. LITTING 207

PAPERS

Emission of negative ions of oxygen from dispenser cathodes. By N. A. SURPLICE 214, 220
Some aspects of the self-heating and ignition of solid cellulosic materials. By P. H. THOMAS and P. C. BOWES 222
Some visco-elastic properties of nylon. By F. L. WARBURTON, J. F. P. JAMES and S. HOTHER-LUSHINGTON 230
Photographic technique for the determination of metal cutting temperatures. By G. BOOTHROYD . . . 238
Ultrasonic relaxation in polyatomic gases and vapours. By S. K. K. JATKAR and D. D. DESHPANDE . . . 243
Simplifying the Fourier integral by adequate specification. By C. G. MAYO and J. W. HEAD 248
Study of electric breakdown of liquid dielectrics using Schlieren optical techniques. By B. FARAZMAND 251
Use of helium flush in the vacuum deposition of thin films. By G. N. SRIVASTAVA and G. D. SCOTT 255
The colour-temperature of an incandescent helical filament. By J. E. GIBBS and G. W. GORDON-SMITH. . 257

NOTES AND NEWS

New books 254
Notes and comments 256, 259

JUNE 1961

CONFERENCE REPORT

Summarized Proceedings of a Conference on Physics of Polymers—Bristol, January 1961. By D. C. BASSETT . 261

PAPERS

Assessment of the accuracy of 'flying-spot' scanning for the measurement of microscopic particles. By C. G. L. FURMIDGE 268
Gaseous diffusion in porous media. Part 3—Wet granular materials. By J. A. CURRIE 275
Electrode contamination and arc formation. By H. DE B. KNIGHT 282
Comprehensive study of the ion pumping of the noble gases. By B. COBIC, G. CARTER and J. H. LECK . . 288
End effects in a falling-sphere viscometer. By A. D. MAUDE 293
Technique for the rapid, accurate and strain-free machining of metallic single crystals. By M. COLE, I. A. BUCKLOW and C. W. B. GRIGSON 296
Use of ring-shaped neutron sources in a research reactor. By P. N. COOPER, K. FIRTH and K. G. STEPHENS . 298
Temperatures under steadily heated floors. By I. G. DONALDSON 300
Small signal impedance of a solid core inductance. By F. KOLLAR and R. D. RUSSELL 307

CORRESPONDENCE

Electrical characteristics of diffused InAs p–n junctions. From G. LUCOVSKY 311

NOTES AND NEWS

New books 306, 310
Notes and comments 312

PAGE

JULY 1961

CONFERENCE REPORT AND ASSOCIATED PAPERS

London symposium on electrical contacts. By M. R. HOPKINS 313
The physics of electrical contact phenomena. By F. LLEWELLYN JONES 318

INVITED ARTICLES

Hydrodynamic research. By F. S. BURT 323
Physics of the ocean. By G. E. R. DEACON 329

PAPERS

Calculation of the behaviour of rubber-covered pressure rollers. By G. J. PARISH 333
Photosensitivity and speed of response in cadmium sulphide. By D. SHAW 337
The influence of an electric field on the growth of copper whiskers. By T. HOFFMANN, J. MAZUR, J. NIKLIBORC and J. RAFAŁOWICZ. 342
Magnetic method for the estimation of ferrite in stainless steel welds. By G. C. CURTIS and J. SHERWIN . . 344
The matching of Pierce guns to tunnels. By C. J. MILNER and K. J. AUSBURN 346

CORRESPONDENCE

Note on the mechanical analogy of a viscoelastic fluid. From P. G. MORGAN 348
Size distribution determinations of non-volatile droplets by light and electron microscopy. From W. J. HARRIS 348
Creep of cylinders in torsion. From R. V. HESKETH . 349

NOTES AND NEWS

New books 332, 341, 351
Notes and comments 351

AUGUST 1961

CONFERENCE REPORT AND ASSOCIATED PAPERS

Summarized Proceedings of a Symposium on Electronic Devices at Helium Temperatures—London, November 1960. By D. H. PARKINSON 353
Recent developments in high speed superconducting devices. By D. R. YOUNG 359
p–n junctions at very low temperatures. By A. K. JONSCHER 363

INVITED ARTICLE

The electricity of precipitation. By J. A. CHALMERS . 372

PAPERS

An induction furnace to attain temperatures above 3000° C in controlled atmospheres. By L. C. F. BLACKMAN, P. H. DUNDAS, A. W. MOORE and A. R. UBBELOHDE . 377
Pumping of argon, nitrogen and hydrogen in a Bayard–Alpert gauge. By B. COBIC, G. CARTER and J. H. LECK 384
Various electrical analogues, incorporating negative resistances, for the solution of problems in elasticity. By S. C. REDSHAW and K. R. RUSHTON . . . 390
Sorption and replacement of ionized noble gases at a tungsten surface. By R. B. BURTT, J. S. COLLIGON and J. H. LECK 396
Emissivity errors of infra-red pyrometers in relation to spectral response. By P. M. REYNOLDS . . . 401
Bandwidth of a moon communication circuit. By J. V. EVANS 406
Production of accurate gas and vapour mixtures. By D. W. HILL 410
Four-probe resistivity measurements on small circular specimens. By D. E. VAUGHAN 414

PAGE

Dielectric properties of ammonium dihydrogen phosphate at very high frequencies. By E. RUSHTON . . . 417
Time lags in the electrical breakdown of glass immersed in water. By M. N. AZAM and H. DICKINSON . . 419
Method of determining the specimen–surface displacement of an x-ray diffractometer. By B. W. DELF . . 421

NOTES AND NEWS

Notes and comments 424

SEPTEMBER 1961

CONFERENCE REPORTS

The Birmingham Conference on Nuclear Physics—April 1961. By W. E. BURCHAM, G. H. BURKHARDT and P. M. ROLPH 425
Summarized Proceedings of a Conference on X-ray Analysis—Glasgow, April 1961. By J. C. SPEAKMAN . 429

INVITED ARTICLE

Thermionic generation of electricity. By M. A. CAYLESS 433

PAPERS

Investigation of the electrical contact properties of granular carbon aggregates. By E. D. MACKLEN . 443
Influence of alloy structure on the transfer of matter. By W. MERL 447
Measurement of the dielectric constant of filament nylon in a transverse field. By F. S. WARD 450
Spectroturbidimetry of emulsions. By J. D. S. GOULDEN 456
Photovoltaic response of selenium barrier layer cells to x-radiation in the energy range of 15–40 kev. By R. FEINBERG and G. E. RHEAD 461
Cathode sputtering in inert-gas glow discharges. By B. J. STOCKER 465

NEW BOOKS SECTION

Reviews and notices 469

PAPERS (*continued*)

Measurement of metal transfer in electrical contacts by the radioactive isotope method. By F. LLEWELLYN JONES, M. R. HOPKINS and C. R. JONES . . . 485
Experimental determination of the wall effect for spheres falling axially in cylindrical vessels. By V. FIDLERIS and R. L. WHITMORE 490
Epitaxy and twinning in foils of some noble metals condensed upon lithium fluoride and mica. By M. J. HALL and M. W. THOMPSON 495
Study of the performance of dielectric thin film beam dividing systems. By L. A. CATALAN and T. PUTNER 499
Counting statistics in x-ray spectroscopy. By R. C. STANLEY 503
Measurement of electrical resistivity by a mutual inductance method. By R. G. CHAMBERS and J. G. PARK . 507
Photoconduction in cadmium sulphide. By J. F. DUNCAN and D. N. SITHARAMA RAO 511
Rheology of granular material II—a method for the determination of the intergranular cohesion. By M. M. BENARIE 514
Spectral response of antimony–caesium photocathodes. By J. B. BIRKS and I. H. MUNRO 519

CORRESPONDENCE

Average kinetic energy of diffusing particles in discharge and electron tubes. From M. A. CAYLESS . . . 523
Dislocation etch pits in polished lead telluride. From G. P. TILLY 524
Electron-optical conditions at Venetian blind type dynodes and their effects in photomultipliers. From P. GÖRLICH, A. KROHS and H.-J. POHL 525

PAGE

Araldite used for model analysis of composite structures. From R. S. ALWAR 526

NOTES AND NEWS

Notes and comments 510, 513, 518, 527

OCTOBER 1961

INVITED ARTICLES

Lead zirconate–titanate piezoelectric ceramics. By A. E. CRAWFORD 529
Magnetism in transition metal alloys. By W. M. LOMER 535
The reactor as a source of radiation. By A. J. SALMON . 537

PAPERS

Principles governing the viscous flow of suspensoid sols. By T. V. STARKEY, V. A. HEWLETT, J. H. A. ROBERTS and R. E. JAMES 545
Diamond knife ultra microtomy of metals and the structure of microtomed sections. By R. PHILLIPS . . 554
Capstan equation for strings with rigidity. By I. M. STUART 559
Contact ferroelectric effect applied to plastic flow. By B. S. SATYANARAYANA 563
Estimation of the total emission of a diode under normal operating conditions without removing the space-charge. By C. S. BULL and R. K. FITCH 566
Hypothesis on the ionic mechanism of the growth of whiskers obtained by reduction of the metal halides. By J. MAZUR and J. RAFAŁOWICZ 569
Experimental estimation of the degree of the ionic thermal dissociation of the CuI vapour in the air atmosphere between 350° c and 550° c. By J. MAZUR and J. RAFAŁOWICZ 572
Optical absorption of some oxides in the Schumann ultraviolet region. By W. P. DOYLE and F. J. P. CLARKE . 574
Improved electron gun for zone-melting bombardment. By M. COLE, C. FISHER and I. A. BUCKLOW . . . 577

CORRESPONDENCE

Electrical conductivity and photoconductivity of thin anthracene layers in vacuum. From C. GHEORGHIȚĂ-OANCEA 579
Annealing of residual stress in silicon monoxide films. From J. PRIEST and H. L. CASWELL; and L. HOLLAND and T. PUTNER 580

NOTES AND NEWS

New books 582
Notes and comments 544, 562, 576, 584

NOVEMBER 1961

CONFERENCE REPORTS AND ASSOCIATED PAPERS

Conference on Electron Microscopy—Nottingham, July 1961. By P. M. KELLY and R. REED 585
Summarized Proceedings of a Conference on Thermoelectricity—Durham, July 1961. By B. R. PAMPLIN . 592
Some problems in the development of a commercial thermoelectric refrigerator. By T. B. BURNETT, H. O. LORCH and J. E. THOMPSON 595

PAPERS

Clean glass substrates for Permalloy films. By A. F. FRAY and S. NIELSEN 603
High power heating of a cylindrical tube by electron bombardment of the inside surface. By R. E. HAIGH and P. H. DAWSON 609

PAGE

Recording micro-densitometer employing beta-radiation from a carbon-14 radioisotope source. By R. A. LLOYD and E. H. ADAMS 614
Light transmission technique for following continuous changes in optical retardation in polymer films. By J. H. MAGILL 618
Dynamic compression measurements of cushioning materials utilizing an analogue computer. By B. H. VENNING and J. GUTTRIDGE 621
Simple viscometer for use with low melting point metals. By D. M. COTTINGHAM 625
Some electromechanical effects on dielectrics. By K. C. KAO 629
Compensating radiometer. By A. R. KING 633
Influence of the electric field on the silver and iron whiskers growth. By T. HOFFMANN, J. MAZUR, J. NIKLIBORC and J. RAFAŁOWICZ 635
Method of measuring the growth of a control specimen of graphite after irradiation. By P. A. E. CROSSE and W. L. SNOWSILL 637

NOTES AND NEWS

New books 624, 632, 639
Notes and comments 640

DECEMBER 1961

CONFERENCE REPORTS AND ASSOCIATED PAPERS

Rutherford Jubilee International Conference—Manchester, September 1961. By W. E. BURCHAM . . 641
Proceedings of the Symposium on Tunnel Diodes—London, February 1961. By I. G. CRESSELL . . 646
Indium antimonide tunnel diodes. By K. F. HULME . 651
The physics of the tunnel diode. By A. K. JONSCHER . 654

INVITED ARTICLE

Electroluminescent devices. By H. K. HENISCH . . 660

PAPERS

Vacuum-deposited metal film resistors. By G. SIDDALL and B. A. PROBYN 668
Measurement of thermal conductivity of semiconductors at high temperatures. By A. D. STUCKES . . . 675

PAGE

Rapid method for determining v.h.f. dielectric parameters for liquids and solutions using standing wave procedures. By M. G. CORFIELD, J. HORZELSKI and A. H. PRICE 680
On the transfer relation of a manometer for liquid level measurements. By S. DUTTA ROY 683
Investigation of a method of growing crystals of GaP and GaAs from the vapour phase. By G. R. ANTELL . 687
Diffusion in heterogeneous media: lattices of parallelepipeds in a continuous phase. By R. M. BARRER and J. H. PETROPOULOS 691
Photoelectric processes in calcium oxide coated cathodes. By C. H. B. MEE and F. A. VICK 698
Microwave appraisal of maser crystals. By J. S. THORP, J. H. PACE and D. F. SAMPSON 705
Determination of the emissivity, for total radiation, of small diameter platinum–10% rhodium wires in the temperature range 600–1450° c. By D. BRADLEY and A. G. ENTWISTLE 708
Observation accuracy in relation to the British Standards Institution's draft recommendation for industrial instrument scale design. By J. SPENCER 712

CORRESPONDENCE

Dielectric behaviour of mercuric oxide. From Y. L. YOUSEF, S. A. ZAHER, Z. HANAFY and T. M. SALEM . 715
Anomalous variation of surface tension. From R. W. COTTERHILL and R. K. FITCH 716
Thermal expansion of bismuth telluride. From K. N. R. TAYLOR 717

NOTES AND NEWS

New books 674, 686, 690, 704, 720
Notes and comments 718

SUBJECT INDEX TO VOLUME 12 721

AUTHOR INDEX TO VOLUME 12 725

NEW BOOKS INDEX TO VOLUME 12 729

INDEX TO VOLUME 12

SUBJECTS

(IA) denotes invited article, and (C) correspondence

Absorption, optical of oxides in the Schumann ultra-violet region 574
Alcohols, polyhydric, and water, periodic discontinuities in mutual potential energy of molecules 65
Alloy structure, influence on transfer of matter 447
Alloys, Pt–10%Rh, *see* Platinum
Alloys, transition metal, magnetism in (IA) 535
Aluminium, 99·7%, spectral emissivity between 200 and 540° c 111
Ammonium dihydrogen phosphate, dielectric properties at very high frequencies 417
Amplification through stimulated emission (IA) 197
Analogue computer, for analysis of stress patterns near holes 178
Analogue computer, use for dynamic compression measurements of cushioning materials 621
Analogues, electrical, incorporating negative resistances, for solution of problems in elasticity 390
Analogy, mechanical of visco-elastic fluid (C) 348
Analysis, magnetic, *see* Magnetic analysis
Annealing of residual stress in SiO films (C) 580
Araldite for stress analysis of models of composite structures (C) 526
Arc formation due to electrode contamination 282

Bandwidth of moon communication circuit 406
Barium films, *see under* Films
Bayard–Alpert gauge for pumping argon, nitrogen and hydrogen 384
Beam dividing systems, dielectric thin film, performance 499
Bearings, journal, shaft position indicator for 73
Beta-radiation from ^{14}C radioisotope source, use in recording micro-densitometer 614
Bismuth telluride, thermal expansion of (C) 716
Bonding layers to flat surfaces for strain measurement 8
Boundary value problems in heat conduction, calculation by convolution integrals for cylindrical cavity and half space 14
Breakdown, electrical of glass in water, time lags in 419
British Standards Institution's draft recommendation for industrial instrument scale design, observation accuracy 711

Cadmium sulphide, photoconduction in 511
Cadmium sulphide photoconductivity and speed of response 337
Capstan equation for strings with rigidity 559
Carbon monoxide sorption by Ba films: Arizumi and Kotani hypothesis 120
Cathode sputtering in inert-gas glow discharges 465
Cathodes, oxide coated calcium, photoelectric processes in 698
Ceramics, piezoelectric, lead zirconate–titanate (IA) 529
Cohesion, intergranular, method for determination 514
Colour temperature of incandescent helical filament 257
Computer elements, physics of (IA) 207
Compression measurements, dynamic, of cushioning materials using analogue computer 621
Conduction, heat, *see* Heat conduction
Conductivity, electrical, *see* Electrical conductivity
Conference reports:
  Electrical contacts, London, April 1961 313
  Electron microscopy, Nottingham, July 1961 585
  Electronic devices at helium temperatures, London, November 1960 353
  Non-Destructive Testing Group and Société Française de Métallurgie, May 1960 125
  Nuclear physics, Birmingham, April 1961 425
  Physics of polymers, Bristol, January 1961 261
Conference reports—*continued*
  Rutherford Jubilee International Conference, Manchester, September 1961 641
  Stress analysis, University College of North Staffordshire, April 1960 3
  Thermoelectricity, Durham, July 1961 592
  Tunnel diodes, London, February 1961 646
  X-ray analysis, Glasgow, April 1961 429
Conformal transformation, *see* Transformation
Contact, electrical, properties of granular carbon aggregates 443
Contacts, electrical, fundamental physical processes 318
Contacts, electrical, metal transfer measurement by radioactive isotope method 485
Contacts, electrical, symposium, London, April 1961 313
Contacts, influence of alloy structure on transfer of matter 447
Contamination of electrodes and arc formation 282
Counting statistics in x-ray spectroscopy 503
Copper staining of Ge p–n–p alloyed junction transistor sections 193
Copper whiskers, *see* Whiskers
Cords and yarns, rubber models of 147
Coupling factor of piezoelectric ceramic disks 188
Creep of cylinders in torsion (C) 349
Crystals, GaP and GaAs, growth from vapour phase 687
Crystals, growth in glass 10
Crystals, maser, microwave appraisal of 705
Crystals, single, magnesium oxide, etching and polishing 44
Crystals, single, metallic, technique for rapid, accurate and strain-free machining 296
CuI vapour in atmosphere between 350 and 550° c experimental estimation of degree of ionic thermal dissociation 572
Currents, induced, frequency dependence 146
Cushioning materials, dynamic compression measurements using analogue computer 621
Cutting, *see* Diamond

Deformation, surface, investigation of 134
Demagnetization of ferromagnetic particles 155
Densitometer, *see* Micro-densitometer
Deposition, vacuum, *see* Vacuum deposition
Design of scales for industrial instruments (IA) 33
Diamond, knife, ultra-microtomy of metals and structure of microtomed sections 554
Dielectric behaviour of mercuric oxide (C) 715
Dielectric constant of filament nylon, measurement in transverse field 450
Dielectric parameters, v.h.f., for liquids and solutions using standing wave procedures, method for determination 679
Dielectric properties of ammonium dihydrogen phosphate at very high frequencies 417
Dielectrics, electromechanical effects on 629
Dielectrics, liquid, electric breakdown study using Schlieren optical techniques 251
Diffractometer, x-ray, method of determining specimen–surface displacement 421
Diffusion, gaseous, in porous media: Part 3—Wet granular materials 275
Diffusion in heterogeneous media: lattices of parallelepipeds in continuous phase 691
Diode, estimation of total emission under normal operating conditions 566
Discharge, glow, *see* Glow discharge
Droplets, non-volatile, size distribution determination by light and electron microscopy (C) 348

Dynodes, Venetian blind type, electron-optical conditions, and effects in photomultipliers (C) 525

Echo, amplitude, ultrasonic, interpretation 25
Echo-ranger, narrow-beam, for fishery and geological investigations 103
Elasticity, dynamic measurement using resonance methods (C) 80
Elasticity, problems solved by electrical analogues 390
Electric breakdown of liquid dielectrics, study using Schlieren optical techniques 251
Electric field, influence on silver and iron whiskers growth 635
Electrical conductivity and photoconductivity of thin anthracene layers in vacuum (C) 579
Electricity of precipitation (IA) 372
Electricity, thermionic generation (IA) 433
Electrode contamination and arc formation 282
Electroluminescent devices (IA) 660
Electrolytic tank model for determination of stress concentrations 184
Electromechanical effects on dielectrics 629
Electron bombardment, high power heating of cylindrical tube by 609
Electron current, matching of Pierce guns to tunnels 346
Electron gun for zone-melting bombardment 577
Electron guns, Pierce 346
Electron microscope, heating of metallic foils in 115
Electron microscopy, conference, Nottingham, July 1961 585
Electron microscopy used to determine size distribution of non-volatile droplets (C) 348
Electron-optical conditions at Venetian blind type dynodes and effects in photomultipliers (C) 525
Electron tubes, average kinetic energy of diffusing particles (C) 523
Electronic devices at helium temperatures, symposium, London, November 1960 353
Emission of negative ions of oxygen from dispenser cathodes: I. Barium oxide in sintered nickel; II. Barium aluminate in sintered tungsten 214, 220
Emission, total, of diode under normal operating conditions 566
Emissivity, for total radiation, of small diameter Pt–10% Rh wires in temperature range 600–1450° c 708
Emissivity, spectral, *see* Spectral emissivity
Emulsions, monodisperse, turbidity, and applications to polydisperse systems 456
Epitaxy and twinning in foils of noble metals condensed on LiF and mica 495
Equations, physical, transformation into dimensionless parametric form 180
Etch pits, dislocation, in polished lead telluride (C) 524
Etching and polishing magnesium oxide single crystals 44
Exhibition discourses:
  Hydrodynamic research (IA) 323
  Physics of the ocean (IA) 329
Expansion, thermal, *see* Thermal expansion (C)

Ferrite in stainless steel welds, estimation by magnetic method 344
Ferroelectric effect, contact, applied to plastic flow 563
Ferromagnetic particles, demagnetization 155
Fibre assemblies, internal friction in 99
Fields outside uniformly magnetized ellipsoidal or spheroidal structures 160
Filament, incandescent helical, colour temperature 257
Filaments, *see also* Nylon
Film resistors, vacuum-deposited metal 668
Films, barium, sorption by CO: Arizumi and Kotani hypothesis 120
Films, lime, of SiO, annealing of residual stress (C) 580
Films, polymer, continuous changes in optical retardation 618
Films, thin, anthracene, photoconductivity in vacuum (C) 579
Films, thin, dielectric, as beam dividing systems, performance 499
Films, thin, germanium and silicon 92
Films, thin, Permalloy, clean glass substrates for 603
Films, thin transparent, thickness measurement by fluorescence 175
Films, thin, use of helium flush in vacuum deposition 255
Films, thin, vacuum-deposited metal, as resistors 668
Films, thin, *see also* Foils
Fire research, radiometer for 633
Fishery investigations, by narrow beam echo-ranger 103
Floors, heated, temperatures under 300
Flow, plastic, contact ferroelectric effect 563
Flow, viscous, of suspensoid sols 545
Fluid, visco-elastic, mechanical analogy (C) 348
Fluorescence method for measurement of thickness of thin transparent films 175
Fluorimeter for measurement of thin (10^{-4} cm) transparent films 175
Foils, metallic, *see* Metallic foils
Foils, noble metal, condensed on LiF and mica, epitaxy and twinning 495
Fourier integral, simplification by adequate specification 248
Frequency dependence of induced currents 146
Friction, internal, in fibre assemblies 99
Frictional properties of lignum vitae 118
Furnace, induction, *see* Induction furnace

Galvanometer effects and their application (IA) 85
Gas discharge tubes, average kinetic energy of diffusing particles (C) 523
Gases, noble, ion pumping 288
Gases, pumping in Bayard–Alpert gauge 384
Gases and vapours, polyatomic, ultrasonic relaxation in 243
Gases and vapours, production of accurate mixtures for calibration purposes 410
Geological investigations by narrow beam echo-ranger 103
Germanium, preparation of thin films 92
Glass, growth of crystals in 10
Glass immersed in water, time lags in electrical breakdown 419
Glass substrates, clean, for Permalloy films 603
Glow discharge, inert-gas, cathode sputtering in 465
Granular carbon aggregates, electrical contact properties 443
Granular material II, rheology: method for determination of intergranular cohesion 514
Granular materials, wet, gaseous diffusion in 275
Granular shape factors measurement and porosity and surface area of bed 172
Graphite, measurement of growth of control specimen after irradiation 637
Growth of crystals of GaP and GaAs from vapour phase 687

Heat conduction, calculation of boundary value problems, for cylindrical cavity and half space 14
Heating, high power, of cylindrical tube by electron bombardment of inside surface 609
Heating of metallic foils in electron microscope 115
Helium flush in vacuum deposition of thin films 255
Helium temperatures, electronic devices at, symposium, London, November 1960 353
Heterogeneous media, diffusion in: lattices of parallelepipeds in continuous phase 691
High temperatures in controlled atmospheres, induction furnace for 377
Hydrodynamic research (Exhibition discourse) (IA) 323

Ignition of cellulosic materials 222
Impedance, small signal, of solid core inductance 307
Indium antimonide tunnel diodes 651
Inductance, solid core, small signal impedance 307
Induction furnace for temperatures above 3000° c in controlled atmospheres 377
Infra-red radiation, measurement of 633

Instrument, industrial, design of scales for (IA) 33
Instrument scale design, industrial, observation accuracy and B.S.I. draft recommendation 712
Insulation of heated floors 300
Integrals, convolution, for calculating boundary value problems in heat conduction 14
Interferometry, measurement of small angular displacements—double-passed Jamin interferometer 20
Internal friction in fibre assemblies 99
Ion bombardment by noble gases, sorption and replacement at tungsten surface 396
Ion pumping of the noble gases 288
Ion sources in mass spectrometers, emission of negative ions of oxygen 214, 220
Ionic mechanism of growth of whiskers obtained by reduction of metal halides 569
Ionic thermal dissociation of CuI vapour in atmosphere between 350 and 550° c 572
Ionization gauge, Bayard–Alpert, for pumping argon, nitrogen and hydrogen 384
Iron whiskers, *see* Whiskers
Irradiation of graphite, measurement of growth after 637
Isotopes, radioactive for measurement of metal transfer in electrical contacts 485

Kerr's constant, pulse method for measuring 50

Lead telluride, polished, dislocation etch pits in (C) 524
Lead zirconate–titanate piezoelectric ceramics (IA) 529
Light transmission technique for following continuous changes in optical retardation in polymer films 618
Lignum vitae, frictional properties 118
Liquid level measurements, transfer relation of manometer for 683
Liquids, cylinder wall effect of spheres falling through, experimental determination 490
Liquids and solutions, determination of v.h.f. dielectric parameters by standing wave procedures 680
Low temperatures, p–n junctions at very low temperatures 363
Lubrication, hydrodynamic, shaft position indicator for use with journal bearings 73

Magnetic analysis, recent advances in 141
Magnetic materials, soft, influence of method of demagnetization on permeability (C) 81
Magnetic method for estimation of ferrite in stainless steel welds 344
Magnetic structures, ellipsoidal or spheroid, external fields 160
Magnetism in transition metal alloys (IA) 535
Manometer for liquid level measurements, transfer relation of 683
Maser (IA) 197
Maser crystals, microwave appraisal of 705
Materials, brittle, measurement of tensile strength (C) 29
Materials, cellulosic, self-heating and ignition 222
Materials, examination by magnetic analysis 141
Materials, fibre insulating board, self-heating and ignition 222
Materials, granular, *see* Granular
Mercuric oxide, dielectric behaviour of (C) 715
Metal cutting, temperatures determination, photographic technique 238
Metal surfaces, smooth spark planing technique for machining 296
Metal transfer in electrical contacts measurement by radioactive isotope method 485
Metal whiskers, *see* Whiskers
Metallic foils in electron microscope, heating 115
Metals, low melting point, measurement of viscosities 625
Method of determining the specimen-surface displacement of x-ray diffractometer 421
Micro-densistometer, recording, employing β-radiation from ^{14}C radioisotope source 614
Microtomy, *see* Ultra-microtomy
Microwave appraisal of maser crystals 705
Moon as reflector in communication systems 406
Mutual inductance method for measurement of electrical resistivity 507

Neutron sources, ring-shaped, use in Merlin reactor 298
Non-destructive measurement of total emission of diode 566
Non-destructive testing, conference report, May 1960 125
Non-destructive testing, structure of 127
Non-destructive testing, *see also* Materials, examination
Nuclear physics, conference, Birmingham, April 1961 425
Nuclear physics, Rutherford Jubilee International Conference, Manchester, September 1961 641
Nylon, filament, measurement of dielectric constant in transverse field 450
Nylon, visco-elastic properties of 230

Observation accuracy and B.S.I.'s draft recommendation 712
Ocean, physics of (Exhibition discourse) (IA) 329
Optical absorption, *see* Absorption, optical
Optical retardation in polymer films, continuous changes in 618
Optical techniques, Schlieren, for study of breakdown of liquid dielectrics 251

Particles, microscopic, accuracy of measurement by 'flying-spot' scanning 268
Particles, spherical, packing and significance of properties of materials 169
Permalloy films, *see* Films, thin Permalloy
Permeability of heterogeneous medium consisting of lattices of parallelepipeds in continuous phase 691
Permeability of soft magnetic materials, influence of method of demagnetization on (C) 81
Permittivities of liquids and solutions in range 250 to 900 Mc/s, measurement using standing wave procedures 680
Photocathodes, Sb–Cs, spectral response 519
Photoconduction in cadmium sulphide 511
Photoconductivity of thin anthracene layers in vacuum (C) 579
Photoelectric processes in calcium oxide coated cathodes 697
Photographic technique for determination of metal cutting temperatures 238
Photomultipliers, effect of electron-optical conditions in Venetian blind type dynodes (C) 525
Photosensitivity and speed of response in cadmium sulphide 337
Photovoltaic response of selenium barrier layer cells to x-radiation in energy range 15–40 kev 461
Physics of the ocean (Exhibition discourse) (IA) 329
Piezoelectric ceramic disks, coupling factor 188
Piezoelectric ceramics, lead zirconate–titanate (IA) 529
Plastic flow, *see* Flow, plastic
Platinum–10% rhodium wires in temperature range 600–1450° c, determination of emissivity for total radiation 708
p–n junctions at very low temperatures 363
p–n junctions, diffused InAs, electrical characteristics (C) 311
Polishing and etching, magnesium oxide single crystals 44
Polymer films, *see* Films, polymer
Polymers, physics of, conference, Bristol, January 1961 261
Porosity and surface area of a granular bed from air flow measurements 172
Porous media, gaseous diffusion in: Part 3—Wet granular materials 275
Precipitation, electricity of (IA) 372
Pressure rollers, rubber-covered, behaviour 333
Pulse method for measuring Kerr's constant 50
Pumping of argon, nitrogen and hydrogen in Bayard–Alpert gauge 384
Pyrometers, infra-red, emissivity errors and spectral response 401

Radiation, infra-red, *see* Infra-red
Radiation, reactor as source of (IA) 537

Radiometer, compensating 633
Reactor as source of radiation (IA) 537
Reactor, Merlin, ring-shaped neutron sources for 298
Reactor, power, heating of test fuel element 609
Refrigerator, thermoelectric, commercial problems in development 595
Relaxation, ultrasonic, *see* Ultrasonic relaxation
Resistances, negative, use in solution of problems in elasticity 390
Resistivity, electrical, measured by mutual inductance method 507
Resistivity measurements on small circular specimens, optimum probe array 414
Resistors, vacuum-deposited metal film 668
Resonance methods for dynamic measurement of elasticity (C) 80
Rheology of granular material II: method for determination of intergranular cohesion 514
Rigidity, string, effect on capstan equation 559
Rubber-covered pressure rollers, behaviour 333
Rubber models of cords and yarns 147
Rutherford Jubilee International Conference, Manchester, September 1961 641

Scales for industrial instruments, design (IA) 33
Scanning, 'flying-spot' accuracy of measurement of microscopic particles 268
Schlieren optical techniques for study of breakdown of liquid dielectrics 251
Selenium barrier layer cells, photovoltaic response to x-radiation in energy range 15–40 kev 461
Semiconductors at high temperatures, measurement of thermal conductivity 675
Semiconductors, InAs p–n junction, diffused, electrical characteristics (C) 311
Semiconductors, *see also* Cadmium sulphide, and p–n junctions
Shaft position indicator for use with journal bearings 73
Shear in suspensoid sols 545
Silicon, preparation of thin films 92
Silver whiskers, *see* Whiskers
Size distribution determinations of non-volatile droplets by light and electron microscopy (C) 348
Soils, *see* Granular materials
Sols, suspensoid, principles governing viscous flow 545
Sorption and replacement of ionized gases at tungsten surface 396
Spark planing for production of flat smooth metal surfaces 296
Spectral emissivity of 99·7% aluminium between 200 and 540° c 111
Spectrophotometers, modifications required for turbidity measurements 456
Spectroscopy, x-ray, *see* X-ray spectroscopy
Spectroturbidimetry of emulsions 456
Spectral response of photocathodes 519
Sputtering, cathode, *see* Cathode sputtering
Stainless steel welds, estimation of ferrite by magnetic method 344
Strain measurement on flat surfaces, photoelastic technique 8
Stress, residual, annealing in SiO films (C) 580
Stress analysis, conference report, University College of North Staffordshire, April 1960 3
Stress analysis, of composite structures, use of Araldite in models (C) 526
Stress concentrations determined with electrolytic tank model 184
Stress patterns near holes, analogue computer for 178
String rigidity, effect on capstan equation 559
Superconducting devices, high speed, recent developments 359
Surface area and porosity of granular bed from air flow measurements 172
Surface deformation, investigation of 134
Surfaces, flat, photoelastic technique for strain measurement 8
Surface tension, anomalous variation of (C) 716

Temperatures, helium, *see* Helium temperatures
Temperatures, high, *see* High temperatures
Temperatures, low, *see* Low temperatures
Temperatures under steadily heated floors 300
Tensile strength of brittle materials, measurement (C) 29
Testing, non-destructive, *see* Non-destructive testing
Thermal conductivity of semiconductors at high temperatures, measurement of 675
Thermal expansion of bismuth telluride (C) 717
Thermionic generation of electricity (IA) 433
Thermocouples, *see* Platinum–rhodium
Thermoelectric refrigerator, commercial, problems in development 595
Thermoelectricity, conference, Durham, July 1961 592
Thin films, *see* Films, thin
Time lags in electrical breakdown strength of glass in water 419
Torsion, creep of cylinders in (C) 349
Transfer relation of manometer for liquid level measurements 683
Transformation, conformal, by analogue computer 178
Transformation of physical equations into dimensionless parametric form 180
Transistors, Ge p–n–p alloyed junction, copper staining of 193
Tunnel diodes, indium antimonide 651
Tunnel diodes, physics of 654
Tunnel diodes, report on symposium, London February 1961 646

Ultra-microtomy of metals and structure of microtomed sections 554
Ultrasonic echo amplitude, interpretation 25
Ultrasonic relaxation in polyatomic gases and vapours 243
Ultra-violet region, Schumann, optical absorption of oxides 574

Vacuum-deposited metal film resistors 668
Vacuum deposition of thin films, use of helium flush 255
Vapours and gases, polyatomic, ultrasonic relaxation in 243
Vapours and gases, production of accurate mixtures for calibration purposes 410
Visco-elastic fluid, mechanical analogy (C) 348
Visco-elastic properties of nylon 230
Viscometer, capillary, of negligible kinetic energy effect 94
Viscometer, falling-sphere, end effects 293
Viscometer for use with low melting point metals 625
Viscosity, periodic discontinuities in mutual potential energy of molecules of water and polyhydric alcohols 65
Viscous flow, *see* Flow, viscous

Wall effect of spheres falling axially in cylindrical vessels, experimental determination 490
Whiskers, growth, copper, influence of electric field 342
Whiskers growth, silver and iron, influence of electric field 635
Whiskers, metal, obtained by reduction of metal halides, hypothesis on ionic mechanism 569
Wire emissivity, Pt–10% Rh, in temperature range 600–1450° c 708
Wood density in trees, measurement of variation by recording micro-densitometer 614

X-irradiation, photovoltaic response of selenium barrier layer cells in energy range 15–40 kev 461
X-ray analysis, conference, Glasgow, April 1961 429
X-ray diffractometer, *see* Diffractometer, x-ray 421
X-ray spectroscopy, counting statistics 503

Yarns and cords, rubber models of 147

Zone-melting bombardment, electron gun 577

AUTHORS

(IA) denotes invited article, and (C) correspondence

Adams, E. H., with Lloyd, R. A.: Recording micro-densitometer employing beta-radiation from carbon-14 radioisotope source 614

Ahsanullah, A. K. M., with Qurashi, M. M.: Investigation of periodic discontinuities in mutual potential energy of molecules of water and some polyhydric alcohols 65

Alwar, R. S.: Araldite used for model analysis of composite structures (C) 526

Amson, J. C.: Method for consistent transformation of physical equations into non-dimensionally parametric form 180

Andrews, K. W.: Joint Conference of Non-Destructive Testing Group and Société Française de Métallurgie 125

Andrews, K. W.: Structure of non-destructive testing 127

Antell, G. R.: Investigation of method of growing crystals of GaP and GaAs from vapour phase 686

Ausburn, K. J., with Milner, C. J.: Matching of Pierce guns to tunnels 346

Azam, M. N., and Dickinson, H.: Time lags in electrical breakdown of glass immersed in water 419

Barrer, R. M., and Petropoulos, J. H.: Diffusion in heterogeneous media: lattices of parallelepipeds in continuous phase 690

Bassett, D. C.: Summarized Proceedings of Conference on Physics of Polymers, Bristol, January 1961 261

Benarie, M. M.: Rheology of granular material II—method for determination of intergranular cohesion 514

Berenbaum, R., with Ormerod, A., and Brodie, I.: Measurement of tensile strength of brittle materials (C) 29

Birks, J. B., and Munro, I. H.: Spectral response of antimony–caesium photocathodes 519

Blackman, L. C. F., Dundas, P. H., Moore, A. W., and Ubbelohde, A. R.: Induction furnace to attain temperatures above 3000° C in controlled atmospheres 377

Boothroyd, G.: Photographic technique for determination of metal cutting temperatures 238

Bowes, P. C., with Thomas, P. H.: Some aspects of self-heating and ignition of solid cellulosic materials 222

Bradfield, G.: Dynamic measurement of elasticity using resonance methods (C) 80

Bradley, D., and Entwistle, A. G.: Determination of emissivity, for total radiation, of small diameter platinum–10% rhodium wires in temperature range 600–1450° C 707

Brodie, I., with Ormerod, A., and Berenbaum, R.: Measurement of tensile strength of brittle materials (C) 29

Bucklow, I. A., with Cole, M., and Fisher, C.: Improved electron gun for zone-melting bombardment 577

Bucklow, I. A., with Cole, M., and Grigson, C. W. B.: Technique for rapid, accurate and strain-free machining of metallic single crystals 296

Bull, C. S., and Fitch, R. K.: Estimation of total emission of diode under normal operating conditions without removing space-charge 566

Bunt, B. P.: Measurement of thickness of thin transparent films using fluorescence 175

Burcham, W. E.: Rutherford Jubilee International Conference, Manchester, September 1961 641

Burcham, W. E., Burkhardt, G. H., and Rolph, P. M.: Birmingham Conference on Nuclear Physics, April 1961 425

Burkhardt, G. H., with Burcham, W. E., and Rolph, P. M.: Birmingham Conference on Nuclear Physics, April 1961 425

Burnett, T. B., Lorch, H. O., and Thompson, J. E.: Some problems in development of commercial thermoelectric refrigerator 595

Burt, F. S.: Hydrodynamic research (IA) 323

Burtt, R. B., Colligon, J. S., and Leck, J. H.: Sorption and replacement of ionized noble gases at tungsten surface 396

Calvert, T. W. G.: Determination of stress concentrations with electrolytic tank model 184

Campbell, D. S., and MacSwan, A. M.: Coupling factor of piezoelectric ceramic disks 188

Carter, G., with Cobic, B., and Leck, J. H.: Comprehensive study of ion pumping of noble gases 288

Carter, G., with Cobic, B., and Leck, J. H.: Pumping of argon, nitrogen and hydrogen in Bayard–Alpert gauge 384

Caswell, H. L., with Priest, J.: Annealing of residual stress in silicon monoxide films (C) 580

Catalan, L. A., and Putner, T.: Study of the performance of dielectric thin film beam dividing systems 499

Cayless, M. A.: Thermionic generation of electricity (IA) 433

Cayless, M. A.: Average kinetic energy of diffusing particles in discharge and electron tubes (C) 523

Caw, W. A., and Wylie, R. G.: Capillary viscometers of negligible kinetic energy effect 94

Chalmers, J. A.: Electricity of precipitation (IA) 372

Chambers, R. G., and Park, J. G.: Measurement of electrical resistivity by mutual inductance method 507

Chang, H.: Fields external to open-structure magnetic devices represented by ellipsoid or spheroid 160

Clarke, F. J. P., with Doyle, W. P.: Optical absorption of some oxides in Schumann ultra-violet region 574

Clarke, F. J. P., with Ghosh, T. K.: Etching and polishing studies on magnesium oxide single crystals 44

Clyne, G., Fessler, H., and Wilson R. W.: Improvements of photoelastic technique for strain measurement on flat surfaces 8

Cobic, B., Carter, G., and Leck, J. H.: Comprehensive study of ion pumping of noble gases 288

Cobic, B., Carter, G., and Leck, J. H.: Pumping of argon, nitrogen and hydrogen in Bayard–Alpert gauge 384

Cole, M., Bucklow, I. A., and Grigson, C. W. B.: Technique for rapid, accurate and strain-free machining of metallic single crystals 296

Cole, M., Fisher, C., and Bucklow, I. A.: Improved electron gun for zone-melting bombardment 577

Colligon, J. S., with Burtt, R. B., and Leck, J. H.: Sorption and replacement of ionized noble gases at tungsten surface 396

Cooper, P. N., Firth, K., and Stephens, K. G.: Use of ring-shaped neutron sources in research reactor 298

Corfield, M. G., Horzelski, J., and Price, A. H.: Rapid method for determining v.h.f. dielectric parameters for liquids and solutions using standing wave procedures 679

Cotterhill, R. W., and Fitch, R. K.: Anomalous variation of surface tension (C) 715

Cottingham, D. M.: Simple viscometer for use with low melting point metals 625

Crawford, A. E.: Lead zirconate–titanate piezoelectric ceramics (IA) 529

Cressell, I. G.: Proceedings of Symposium on Tunnel Diodes, London, February 1961 646

Crosse, P. A. E., and Snowsill, W. L.: Method of measuring growth of control specimen of graphite after irradiation 637

Currie, J. A.: Gaseous diffusion in porous media. Part 3—Wet granular materials 275

Curtis, G. C., and Sherwin, J.: Magnetic method for estimation of ferrite in stainless steel welds 344

Dawson, P. H., with Haigh, R. E.: High power heating of cylindrical tube by electron bombardment of inside surface 609

Deacon, G. E. R.: Physics of the ocean (IA) 329

Delf, B. W.: Method of determining specimen-surface displacement of x-ray diffractometer 421

Deshpande, D. D., with Jatkar, S. K. K.: Ultrasonic relaxation in polyatomic gases and vapours 243

Dickinson, H., with Azam, M. N.: Time lags in electrical breakdown of glass immersed in water 419
Donaldson, I. G.: Temperatures under steadily heated floors 300
Doyle, W. P., and Clarke, F. J. P.: Optical absorption of some oxides in Schumann ultra-violet region 574
Duncan, J. F., and Sitharama Rao, D. N.: Photoconduction in cadmium sulphide 511
Dundas, P. H., with Blackman, L. C. F., Moore, A. W., and Ubbelohde, A. R.: Induction furnace to attain temperatures above 3000° c in controlled atmospheres 377
Dutta Roy, S.: On transfer relation of manometer for liquid level measurements 682

Entwistle, A. G., with Bradley, D.: Determination of emissivity, for total radiation, of small diameter platinum–10% rhodium wires in temperature range 600–1450° c 707
Evans, I.: Dynamic measurement of elasticity using resonance methods (C) 80
Evans, J. V.: Bandwidth of a moon communication circuit 406

Farazmand, B.: Study of electric breakdown of liquid dielectrics using Schlieren optical techniques 251
Feinberg, R., and Rhead, G. E.: Photovoltaic response of selenium barrier layer cells to x-radiation in energy range of 15–40 kev 461
Fessler, H., with Clyne, G., and Wilson, R. W.: Improvements of photoelastic technique for strain measurement on flat surfaces 8
Fidleris, V., and Whitmore, R. L.: Experimental determination of wall effect for spheres falling axially in cylindrical vessels 490
Firth, K., with Cooper, P. N., and Stephens, K. G.: Use of ring-shaped neutron sources in research reactor 298
Fisher, C., with Cole, M., and Bucklow, I. A.: Improved electron gun for zone-melting bombardment 577
Fitch, R. K., with Bull, C. S.: Estimation of total emission of diode under normal operating conditions without removing space-charge 566
Fitch, R. K., with Cotterhill, R. W.: Anomalous variation of surface tension (C) 715
Fray, A. F., and Nielsen, S.: Clean glass substrates for Permalloy films 603
Furmidge, C. G. L.; Assessment of accuracy of 'flying-spot' scanning for measurement of microscopic particles 268

Gale, B., and Hale, K. F.: Heating of metallic foils in electron microscope 115
Gheorghiță-Oancea, C.: Electrical conductivity and photoconductivity of thin anthracene layers in vacuum (C) 579
Ghosh, T. K., and Clarke, F. J. P.: Etching and polishing studies on magnesium oxide single crystals 44
Gibbs, J. E., and Gordon-Smith, G. W.: Colour-temperature of incandescent helical filament 257
Giorgi, T. A. G., and Origlio, S.: Barium–carbon monoxide system: hypothesis of Arizumi and Kotani 120
Gordon-Smith, G. W., with Gibbs, J. E.: Colour-temperature of incandescent helical filament 257
Gorlich, P., Krohs, A., and Pohl, H.-J.: Electron-optical conditions at Venetian blind type dynodes and their effects in photomultipliers (C) 525
Goulden, J. D. S.: Spectroturbidimetry of emulsions 456
Graneau, P.: Frequency dependence of induced currents (abstract) 146
Gray, W. A., with Macrae, J. C.: Significance of the properties of materials in the packing of real spherical particles 164
Grigson, C. W. B., with Cole, M., and Bucklow, I. A.: Technique for rapid, accurate and strain-free machining of metallic single crystals 296
Grunberg, L., Scott, D., and Wright, K. H. R.: Investigation of surface deformation 134
Guttridge, J., with Venning, B. H.: Dynamic compression measurements of cushioning materials utilizing analogue computer 621

Haigh, R. E., and Dawson, P. H.: High power heating of cylindrical tube by electron bombardment of inside surface 609
Hale, K. F., with Gale, B.: Heating of metallic foils in electron microscope 115
Hall, M. J., and Thompson, M. W.: Epitaxy and twinning in foils of some noble metals condensed upon lithium fluoride and mica 495
Hanafy, Z., with Yousef, Y. L., Zaher, S. A., and Salam, T. M.: Dielectric behaviour of mercury oxide (C) 714
Hariharan, P., and Sen, D.: Interferometric measurement of small angular displacements Part 2. Double-passed Jamin interferometer 20
Harris, W. J.: Size distribution determinations of non-volatile droplets by light and electron microscopy (C) 348
Head, J. W., with Mayo, C. G.: Simplifying Fourier integral by adequate specification 248
Henisch, H. K.: Electroluminescent devices (IA) 660
Hesketh, R. V.: Creep of cylinders in torsion (C) 349
Hewlett, V. A., with Starkey, T. V., Roberts, J. H. A., and James, R. E.: Principles governing viscous flow of suspensoid sols 545
Hill, D. W.: Production of accurate gas and vapour mixtures 410
Hilsum, C.: Galvanomagnetic effects and their application (IA) 85
Hoffmann, T., Mazur, J., Nikliborc, J., and Rafałowicz, J.: Influence of electric field on growth of copper whiskers 342
Hoffmann, T., Mazur, J., Nikliborc, J., and Rafałowicz, J.: Influence of electric field on silver and iron whiskers growth 635
Holland, L., and Putner, T.: Annealing of residual stress in silicon monoxide films (C) 580
Hopkins, M. R.: London Symposium on Electrical Contacts 313
Hopkins, M. R., with Llewellyn Jones, F., and Jones, C. R.: Measurement of metal transfer in electrical contacts by radioactive isotope method 485
Horżelski, J., with Corfield, M. G., and Price, A. H.: Rapid method for determining v.h.f. dielectric parameters for liquids and solutions using standing wave procedures 679
Hoselitz, K.: Recent advances in magnetic analysis 141
Hother-Lushington, S., with Warburton, F. L., and James, J. F. P.: Some visco-elastic properties of nylon 230
Huffington, J. D.: Internal friction in fibre assemblies 99
Hughes, C. J., with Hunter, J. J.: Shaft position indicator for use with journal bearings 73
Hughes, J. E., with Ratcliffe, S.: Copper staining of p–n–p alloyed junction transistor sections 193
Hughes, J. W., Lewis, I. E., and Wilson, A. J. C.: Comments on A.S.T.M. Powder Data File—6 (C) 194
Hulme, K. F.: Indium antimonide tunnel diodes 651
Hunter, J. J., and Hughes, C. J.: Shaft position indicator for use with journal bearings 73

Irving, B. A.: Preparation of thin films of germanium and silicon 92

James, J. F. P., with Warburton, F. L., and Hother-Lushington, S.: Some visco-elastic properties of nylon 230
James, R. E., with Starkey, T. V., Hewlett, V. A., and Roberts, J. H. A.: Principles governing viscous flow of suspensoid sols 545
Jatkar, S. K. K., and Deshpande, D. D.: Ultrasonic relaxation in polyatomic gases and vapours 243
Jones, C. R., with Llewellyn Jones, F., and Hopkins, M. R.: Measurement of metal transfer in electrical contacts by radioactive isotope method 485
Jones, W. M., and Mills, D.: Porosity and surface area of granular bed from measurements of flow of air through bed and measurements of granular shape factors 172
Jonscher, A. K.: Physics of tunnel diode 654
Jonscher, A. K.: p–n junctions at very low temperatures 363
Jordan, D. W.: Method for calculating boundary value problems in heat conduction for cylindrical cavity and half space, by means of convolution integrals 14

Joss, E. J., and Ross, D. S.: Conformal transformations by analogue computer 178

Kao, K. C.: Some electromechanical effects on dielectrics 629
Kelly, P. M., and Reed, R.: Conference on Electron Microscopy, Nottingham, July 1961 585
King, A. R.: Compensating radiometer 633
Knight, H. De B.: Electrode contamination and arc formation 282
Kollar, F., and Russell, R. D.: Small signal impedance of solid core inductance 307
Konopka, R., with Piekara, A.: Pulse method for measuring Kerr's constant 50
Krohs, A., with Gorlich, P., and Pohl, H.-J.: Electron-optical conditions at Venetian blind type dynodes and their effects in photomultipliers (C) 525

Leck, J. H., with Burtt, R. B., and Colligon, J. S.: Sorption and replacement of ionized noble gases at tungsten surface 396
Leck, J. H., with Cobic, B., and Carter, G.: Comprehensive study of ion pumping of noble gases 288
Leck, J. H., with Cobic, B., and Carter, G.: Pumping of argon, nitrogen and hydrogen in Bayard–Alpert gauge 384
Lewis, I. E., with Hughes, J. W., and Wilson, A. J. C.: Comments on A.S.T.M. Powder Data File—6 (C) 194
Litting, C. N. W.: Physics of computer elements (IA) 207
Llewellyn Jones, F.: Physics of electrical contact phenomena 318
Llewellyn Jones, F., Hopkins, M. R., and Jones, C. R.: Measurement of metal transfer in electrical contacts by radioactive isotope method 485
Lloyd, R. A., and Adams, E. H.: Recording micro-densitometer employing beta-radiation from carbon-14 radioisotope source 614
Lomer, W. M.: Magnetism in transition metal alloys (IA) 535
Lorch, H. O., with Burnett, T. B., and Thompson, J. E.: Some problems in development of commercial thermoelectric refrigerator 595
Lucovsky, G.: Electrical characteristics of diffused InAs p–n junctions (C) 311

Macklen, E. D.: Investigation of electrical contact properties of granular carbon aggregates 443
McLaren, K. G., and Tabor, D.: Frictional properties of lignum vitae 118
Macrae, J. C., and Gray, W. A.: Significance of the properties of materials in packing of real spherical particles 164
MacSwan, A. M., with Campbell, D. S.: Coupling factor of piezoelectric ceramic disks 188
Maddock, A. J.: Design of scales for industrial instruments (IA) 33
Magill, J. H.: Light transmission technique for following continuous changes in optical retardation in polymer films 618
Maude, A. D.: End effects in falling-sphere viscometer 293
Mayo, C. G., and Head, J. W.: Simplifying Fourier integral by adequate specification 248
Mazur, J., with Hoffmann, T., Nikliborc, J., and Rafałowicz, J.: Influence of electric field on growth of copper whiskers 342
Mazur, J., with Hoffmann, T., Nikliborc, J., and Rafałowicz, J.: Influence of electric field on silver and iron whiskers growth 635
Mazur, J., and Rafałowicz, J.: Hypothesis on ionic mechanism of growth of whiskers obtained by reduction of metal halides 569
Mazur, J., and Rafałowicz, J.: Experimental estimation of degree of ionic thermal dissociation of CuI vapour in air atmosphere between 350° c and 550° c 572
Mee, C. H. B., and Vick, F. A.: Photoelectric processes in calcium oxide coated cathodes 697
Merl, W.: Influence of alloy structure on transfer of matter 447
Mills, D., with Jones, W. M.: Porosity and surface area of granular bed from measurements of flow of air through bed and measurements of granular shape factors 172
Milner, C. J., and Ausburn, K. J.: Matching of Pierce guns to tunnels 346
Morgan, P. G.: Note on mechanical analogy of a viscoelastic fluid (C) 348
Morley, J. G.: Growth of crystals in glass 10
Moore, A. W., with Blackman, L. C. F., Dundas, P. H., and Ubbelohde, A. R.: Induction furnace to attain temperatures above 3000° c in controlled atmospheres 377
Munro, I. H., with Birks, J. B.: Spectral response of antimony–caesium photocathodes 519

Nielsen, S., with Fray, A. F.: Clean glass substrates for Permalloy films 603
Nikliborc, J., with Hoffmann, T., Mazur, J., and Rafałowicz, J.: Influence of electric field on growth of copper whiskers 342
Nikliborc, J., with Hoffmann, T., Mazur, J., and Rafałowicz, J.: Influence of electric field on silver and iron whiskers growth 635

Origlio, S., with Giorgi, T. A. G.: Barium–carbon monoxide systems: hypothesis of Arizumi and Kotani 120
Ormerod, A., Berenbaum, R., and Brodie, I.: Measurement of tensile strength of brittle materials (C) 29

Pace, J. H., with Thorp, J. S., and Sampson, D. F.: Microwave appraisal of maser crystals 704
Pamplin, B. R.: Summarized Proceedings of Conference on Thermoelectricity, Durham, July 1961 592
Parish, G. J.: Calculation of behaviour of rubber-covered pressure rollers 333
Park, J. G., with Chambers, R. G.: Measurement of electrical resistivity by mutual inductance method 507
Parkinson, D. H.: Summarized Proceedings of a Symposium on Electronic Devices at Helium Temperatures, London, November 1960 353
Petropoulos, J. H., with Barrer, R. M.: Diffusion in heterogeneous media: lattices of parallelepipeds in continuous phase 690
Phillips, R.: Diamond knife ultra microtomy of metals and structure of microtomed sections 554
Piekara, A., and Konopka, R.: Pulse method for measuring Kerr's constant 50
Pohl, H.-J., with Gorlich, P., and Krohs, A.: Electron-optical conditions at Venetian blind type dynodes and their effects in photomultipliers (C) 525
Pomeroy, C. D.: Summarized Proceedings of Conference on Stress Analysis, University College of North Staffordshire, April 1960 3
Price, A. H., with Corfield, M. G., and Horzelski, J.: Rapid method for determining v.h.f. dielectric parameters for liquids and solutions using standing wave procedures 679
Priest, J., and Caswell, H. L.: Annealing of residual stress in silicon monoxide films (C) 580
Probyn, B. A., with Siddall, G.: Vacuum-deposited metal film resistors 668
Putner, T., with Catalan, L. A.: Study of performance of dielectric thin film beam dividing systems 499
Putner, T., with Holland, L.: Annealing of residual stress in silicon monoxide films (C) 580

Qurashi, M. M., and Ahsanullah, A. K. M.: Investigation of periodic discontinuities in mutual potential energy of molecules of water and some polyhydric alcohols 65

Rafałowicz, J., with Hoffmann, T., Mazur, J., and Nikliborc, J.: Influence of electric field on growth of copper whiskers 342
Rafałowicz, J., with Hoffmann, T., Mazur, J., and Nikliborc, J.: Influence of electric field on silver and iron whiskers growth 635
Rafałowicz, J., with Mazur, J.: Hypothesis on ionic mechanism of growth of whiskers obtained by reduction of metal halides 569
Rafałowicz, J., with Mazur, J.: Experimental estimation of degree of ionic thermal dissociation of CuI vapour in air atmosphere between 350° c and 550° c 572
Ratcliffe, S., and Hughes, J. E.: Copper staining of p–n–p alloyed junction transistor sections 193

Redshaw, S. C., and Rushton, K. R.: Various electrical analogues, incorporating negative resistances, for solution of problems in elasticity 390
Reed, R., with Kelly, P. M.: Conference on Electron Microscopy, Nottingham, July 1961 585
Reynolds, P. M.: Spectral emissivity of 99·7% aluminium between 200 and 540° c 111
Reynolds, P. M.: Emissivity errors of infra-red pyrometers in relation to spectral response 401
Rhead, G. E., with Feinberg, R.: Photovoltaic response of selenium barrier layer cells to x-radiation in energy range of 15–40 kev 461
Roberts, J. H. A., with Starkey, T. V., Hewlett, V. A., and James, R. E.: Principles governing viscous flow of suspensoid sols 545
Rolph, P. M., with Burcham, W. E., and Burkhardt, G. H.: Birmingham Conference on Nuclear Physics, April 1961 425
Ross, D. S., with Joss, E. J.: Conformal transformations by analogue computer 178
Rumsey, J. C. V.: Interpretation of ultrasonic echo amplitude 25
Rushton, E.: Dielectric properties of ammonium dihydrogen phosphate at very high frequencies 417
Rushton, K. R., with Redshaw, S. C.: Various electrical analogues, incorporating negative resistances, for solution of problems in elasticity 390
Russell, R. D., with Kollar, F.: Small signal impedance of solid core inductance 307

Salam, T. M., with Yousef, Y. L., Zaher, S. A., and Hanafy, Z.: Dielectric behaviour of mercury oxide (C) 714
Salmon, A. J.: Reactor as source of radiation (IA) 537
Sampson, D. F., with Thorp, J. S., and Pace, J. H.: Microwave appraisal of maser crystals 704
Satyanarayana, B. S.: Contact ferroelectric effect applied to plastic flow 563
Scott, D., with Grunberg, L., and Wright, K. H. R.: Investigation of surface deformation 134
Scott, G. D., with Srivastava, G. N.: Use of helium flush in vacuum deposition of thin films 255
Sen, D., with Hariharan, P.: Interferometric measurement of small angular displacements Part 2. Double-passed Jamin interferometer 20
Shaw, D.: Photosensitivity and speed of response in cadmium sulphide 337
Sherwin, J., with Curtis, G. C.: Magnetic method for estimation of ferrite in stainless steel welds 344
Siddall, G., and Probyn, B. A.: Vacuum-deposited metal film resistors 668
Sitharama Rao, D. N., with Duncan, J. F.: Photoconduction in cadmium sulphide 511
Smith, F.: Demagnetization of ferromagnetic particles 155
Smith, R. A.: Amplification through stimulated emission—the MASER (IA) 197
Smolinski, A.: Influence of method of demagnetization on permeability in soft magnetic materials (C) 81
Snowsill, W. L., with Crosse, P. A. E.: Method of measuring growth of control specimen of graphite after irradiation 637
Speakman, J. C.: Summarized Proceedings of a Conference on X-ray Analysis, Glasgow, April 1961 429
Spencer, J.: Observation accuracy in relation to British Standards Institution's draft recommendation for industrial instrument scale design 711
Srivastava, G. N., and Scott, G. D.: Use of helium flush in vacuum deposition of thin films 255
Stanley, R. C.: Counting statistics in x-ray spectroscopy 503
Starkey, T. V., Hewlett, V. A., Roberts, J. H. A., and James, R. E.: Principles governing viscous flow of suspensoid sols 545
Stephens, K. G., with Cooper, P. N., and Firth, K.: Use of ring-shaped neutron sources in research reactor 298
Stocker, B. J.: Cathode sputtering in inert-gas glow discharges 465
Stuart, I. M.: Capstan equation for strings with rigidity 559
Stubbs, A. R., with Tucker, M. J.: Narrow-beam echo-ranger for fishery and geological investigations 103
Stuckes, A. K.: Measurement of thermal conductibility of semiconductors at high temperatures 674
Surplice, N. A.: Emission of negative ions of oxygen from dispenser cathodes 214, 220

Tabor, D., with McLaren, K. G.: Frictional properties of lignum vitae 118
Taylor, K. N. R.: Thermal expansion of bismuth telluride (C) 716
Thomas, P. H., and Bowes, P. C.: Some aspects of self-heating and ignition of solid cellulosic materials 222
Thompson, J. E., with Burnett, T. B., and Lorch, H. O.: Some problems in development of commercial thermoelectric refrigerator 595
Thompson, M. W., with Hall, M. J.: Epitaxy and twinning in foils of some noble metals condensed upon lithium fluoride and mica 495
Thorp, J. S., Pace, J. H., and Sampson, D. F.: Microwave appraisal of maser crystals 704
Tilly, G. P.: Dislocation etch pits in polished lead telluride (C) 524
Treloar, L. R. G., with Wilson, N.: Rubber models of yarns and cords; 'doubling' of single rods 147
Tucker, M. J., and Stubbs, A. R.: Narrow-beam echo-ranger for fishery and geological investigations 103

Ubbelohde, A. R., with Blackman, L. C. F., Dundas, P. H., and Moore, A. W.: Induction furnace to attain temperatures above 3000° c in controlled atmospheres 377

Vaughan, D. E.: Four-probe resistivity measurements on small circular specimens 414
Venning, B. H., and Guttridge, J.: Dynamic compression measurements of cushioning materials utilizing analogue computer 621
Vick, F. A., with Mee, C. H. B.: Photoelectric processes in calcium oxide coated cathodes 697

Warburton, F. L., James, J. F. P., and Hother-Lushington, S.: Some visco-elastic properties of nylon 230
Ward, F. S.: Measurement of dielectric constant of filament nylon in transverse field 450
Whitmore, R. L., with Fidleris, V.: Experimental determination of the wall effect for spheres falling axially in cylindrical vessels 490
Wilson, A. J. C., with Hughes, J. W., and Lewis, I. E.: Comments on A.S.T.M. Powder Data File—6 (C) 194
Wilson, N., and Treloar, L. R. G.: Rubber models of yarns and cords; 'doubling' of single rods 147
Wilson, R. W., with Clyne, G., and Fessler, H.: Improvements of photoelastic technique for strain measurement on flat surfaces 8
Wright, K. H. R., with Grunberg, L., and Scott, D.: Investigation of surface deformation 134
Wylie, R. G., with Caw, W. A.: Capillary viscometers of negligible kinetic energy effect 94

Young, D. R.: Recent developments in high speed superconducting devices 359
Yousef, Y. L., Zaher, S. A., Hanafy, Z., and Salam, T. M.: Dielectric behaviour of mercuric oxide (C) 714

Zaher, S. A., with Yousef, Y. L., Hanafy, Z., and Salam, T. M.: Dielectric behaviour of mercury oxide (C) 714

BOOKS

Ajzenberg-Selove, F. (Ed.): *Pure and applied physics* Vol. 9. *Nuclear spectroscopy* Parts A and B 61

Allen, W. D.: *Neutron detection* 60

Aller, L. H.: *The abundance of the elements* 719

d'Azzo, J. J., and Houpis, C. H.: *Feedback control system analysis and synthesis* 58

Baghdady, E. J. (Ed.): *Lectures on communication system theory* 472

Bahn, G. S., and Zukoski, E. E. (Eds): *Kinetics, equilibria and performance of high temperature systems—Proceedings of the First Conference Western States Section, The Combustion Institute* 1959 473

Baker, R. H.: *An introduction to astronomy* 583

Bell, R. P.: *The proton in chemistry* 55

Bergen, J. T. (Ed.): *Viscoelasticity: phenomenological aspects* 63

Berl, W. G. (Ed.): *Physical methods in chemical analysis* Vol. 1 2nd revised edn 470

Bickley, W. G., and Talbot, A.: *An introduction to the theory of vibrating systems* 475

Bishop, R. E. D., and Johnson, D. C.: *The mechanics of vibration* 474

Blanc, D., and Ambrosino, G.: *Eléments de physique nucléaire* 481

Bland, D. R.: *The theory of linear viscoelasticity* 56

Brailsford, F.: *Magnetic materials* 351

Burrows, G.: *Molecular distillation* 484

Buzzati-Traverso, A. A.: *Immediate and low level effects of ionizing radiations* 64

Carlin, B.: *Ultrasonics*, 2nd edn 477

Centre d'information du Cobalt: *Cobalt monograph* 583

Clark, G. L. (Ed.): *The encyclopaedia of spectroscopy*

Clauser, F. H. (Ed.): *Plasma dynamics* 63

Conn, G. K. T. (Ed.): *Semiconductor technology* 310

Conn, G. K. T., and Avery, D. G.: *Infra-red methods: principles and applications* 473

Cosslett, V. E., and Nixon, W. C.: *X-ray microscopy. Cambridge monographs on physics* 477

Crouthamel, C. E. (Ed.): *Applied gamma-ray spectrometry* (*International series of monographs on analytical chemistry* Vol. 2) 54

Department for Scientific and Industrial Research: *Interferometry—National Physical Laboratory Symposium No.* 11 474

Derby, T. K., and Williams, T. I.: *A short history of technology* 472

Désirant, M., and Michiels, J. L. (Eds): *Electromagnetic wave propagation* 64

Désirant, M., and Michiels, J. L. (Eds): *Solid state physics in electronics and telecommunications.* Vols. 1, 2 and 3 63

D'eye, R. W. M., and Wait, E.: *X-ray powder photography in inorganic chemistry* 57

Dockx, S.: *Théorie fondamentale du systeme périodique des éléments* 56

Drummond, J. E.: *Plasma physics* 483

Eck, B.: *Technische Strömungslehre* 583

Eder, F. X.: *Moderne Messmethoden der Physik: Teil* 1 *Mechanik und Akustik* 306

Egli, P. H. (Ed.): *Thermoelectricity* 473

Engel, H., and Thielmann, K. O.: *Kernenergie Technik* 479

Ernst, D.: *Elektronische analogrechner* 58

Eterman, I. I.: *Analogue computers* 470

Faires, R. A., and Parks, B. H.: *Radioisotope laboratory techniques*, 2nd edn 481

Ferri, A., Küchemann, D., and Sterne, L. H. G. (Eds): *Progress in aeronautical sciences*, Vol. 1

Fizichesky Entsiklopedichesky Slovar. Vol. 1. A-D 583

Fréchette, V. D. (Ed.): *Non-crystalline solids* 582

Frenkiel, F. N., and Sears, W. R. (Eds): *Magneto-fluid dynamics* 583

Geilkman, B. T., translated from the Russian: *The statistical theory of phase transitions* 64

Gibson, A. F., Kröger, F. A., and Burgess, R. E. (Eds): *Progress in semiconductors.* Vol. 5 478

Glasstone, S., and Lovberg, R. H.: *Controlled thermonuclear reactions—an introduction to theory and experiment* 480

Goldsmith, W.: *Impact: the theory and physical behaviour of colliding solids* 476

Green, A. E., and Adkins, J. E.: *Large elastic deformations and non-linear continuum mechanics* 475

Greiner, R. A.: *Semiconductor devices and applications* 719

Gruhle, W.: *Elektronische Hilfsmittel des Physikers* 254

Harrison, T. R.: *Radiation pyrometry and its underlying principles of radiant heat transfer* 60

Härtel, W., Degenhart, J., Kübler, A., Sörensen, C., and Tröger, J.: *Lichtstrahl-Oszillographen* 471

Hearmon, R. F. S.: *An introduction to applied anisotropic elasticity* 475

Heidt, L. J., Livingston, R. S., Rabinowitch, E., and Daniels, F.: *Photochemistry in the liquid and solid states* 63

Herdan, G.: *Small particle statistics*, 2nd edn 481

Hertel, H.: *Leichtbau. Bauelemente, Bemessungen und Konstruktionen von Flugzeugen und anderen Leichtbauwerken* 63

Hogarth, C. A., and Blitz, J. (Eds): *Techniques of non-destructive testing* 476

International Atomic Energy Agency Joint Commission on Applied Radioactivity: *Proceedings of the radioactivation analysis symposium held in Vienna in* 1959 306

Kaye, J., and Welsh, J. A. (Eds): *Direct conversion of heat to electricity* 59

Keil, A.: *Werkstoffe für elektrische Kontakte* 471

Knapman, C. E. H. (Ed.): *Gas chromatography abstracts* 64

Knight, H. de B.: *The arc discharge. Its application to power control* 57

Korn, G. A., and Korn, T. M.: *Mathematical handbook for scientists and engineers* 583

Lawson, D. F.: *The technique of photomicrography* 62

Lebedev, S. A. (Ed.): *Computer engineering* 58

Levinger, J. S.: *Nuclear photo-disintegration* 62

McGee, J. D., and Wilcock, W. L. (Eds): *Advances in electronics and electron physics* Vol. XII, *Photo-electronic image devices* 57

McGonnagle, W. J.: *Non-destructive testing* 624

McIntosh, A. B., and Heal, T. J. (Eds): *Materials for nuclear engineers* 480

Marion, J. B., and Fowler, J. L. (Eds): *Fast neutron physics: Part* 1. *Techniques.* Interscience monographs and texts in physics and astronomy Vol. IV 479

Martinot Lagarde, A.: *Similitude physique—exemples d'applications à la méchanique des fluides* 582

Marton, L. (Ed.): *Advances in electronics and electron physics* Vol. XI 57

Mason, D. M., Reynolds, W. C., and Vincenti, W. G. (Eds): *Proceedings of the* 1960 *Heat Transfer and Fluid Mechanics Institute* 306

Meerov, M. V.: *Introduction to the dynamics of automatic regulating of electrical machines* 632

Meissner, C. R. (Ed.): *Vacuum technology transactions. Proceedings of the Sixth National Symposium* 483

Mendelssohn, K. (Ed.): *Progress in cryogenics* Vol. 3 582

Menzel, D. H. (Ed.): *The radio noise spectrum* 481

Miller, J., Gerhauser, J. M., and Matsen, F. A.: *Quantum chemistry integrals and tables* 60

Minkoff, G. J.: *Frozen free radicals* 55

Münzinger, F.: *Atomkraft*, 3rd edn 62

Nicholls, C. M. (Ed.): *Progress in nuclear energy, series IV. Technology, engineering and safety.* Vol. III 478

Pauling, L.: *The nature of the chemical bond and the structure of molecules and crystals*, 3rd edn 56

Payne, A. R., and Scott, J. R.: *Engineering design with rubber* 54

Pease, D. C.: *Histological techniques for electron microscopy* 341

Pirani, M. (Ed.): *Elektrothermie* 473

Powell, C. F., Fowler, P. H., and Perkins, D. H.: *The study of elementary particles by the photographic method* 61

Predvoditelev, A. S.: *Physical gas dynamics* 583
Ralston, O. C.: *Elsevier monographs: Electrostatic separation of mixed granular solids* 471
Redwood, M. R.: *Mechanical waveguides* 476
Ribauld, G.: *Conduction de la chaleur en régime variable* 59
Richards, J. A., Sears, F. W., Wehr, M. R., and Zemansky, M. W.: *Modern university physics* 254
Ridenour, L. N., and Nierenburg, W. A. (Eds): *Modern physics for the engineer* 583
Rider, J. F., and Uslan, S. D.: *Encyclopaedia on cathode-ray oscilloscopes and their uses* 58
Royal College of Science and Technology, Glascow: *Nuclear reactor containment buildings and pressure vessels*—Proceedings of a symposium organized by the Department of Mechanical, Civil and Chemical Engineering of the Royal College of Science and Technology, Glasgow 583
Ruzicka, J. E. (Ed.): *Structural damping* 63
Sangren, W. C.: *Digital computers and nuclear reactor calculations* 332
Schlicke, H. M.: *Essentials of dielectromagnetic engineering* 582
Segrè, E., Hanna, G. C., Deutsch, M., Kofoed-Hansen, O., and McMillan, E. M. (Eds): *Experimental nuclear physics* Vol. 3 61
Seirfert, H. S. (Ed.): *Space technology* 62
Semenov, N. N., translated from the Russian by J. E. S. Bradley: *Some problems of chemical kinetics and reactivity* Vol. 2 56
Siegbahn, K. (Ed.): *Nuclear instruments and methods* 583
Smyth, C. N. (Ed.): *Medical electronics* 63
Stanyukovich, K. P.: *Unsteady motion of continuous media* 56
State Scientific Publishing House; *Fizichesky Entsiklopedichesky Slovar* Vol. I A^{ei} D 583.
Stephenson, R. J.: *Mechanics and properties of matter*, 2nd edn 474
Strutt, M., and Vilbig, F.: *Fortschritte der Hochfrequenztechnik*, Band 5 718
Tatarski, V. I. (Translated by R. A. Silverman): *Wave propagation in a turbulent medium* 483
Taylor, A.: *X-ray metallography* 639
Thom, A., Apelt, C. J., and Temple, G. F. J.: *Field computations in engineering and physics* 482
Thring, M. W. (Ed.): *Nuclear propulsion* 480
Tipton, C. R.: *Reactor handbook* Vol. 1. *Materials*, 2nd revised edn 60
Tschauner, J.: *Einführung in die Theorie der Abtastsysteme* 483
Ubbelohde, A. R., and Lewis, F. A.: *Graphite and its crystal compounds* 55
University of Buffalo: *Carbon: Proceedings of the Fourth Conference held at the University of Buffalo, New York* 310
Wait, J. R. (Ed.): *Overvoltage research and geophysical applications* 59
Zenneck, J., Strutt, M., and Vilbig, F. (Eds): *Fortschritte der Hochfrequenztechnik* 310

PRINTED IN GREAT BRITAIN BY UNWIN BROTHERS LIMITED, WOKING AND LONDON

BRITISH JOURNAL OF

Applied Physics

VOLUME 12 JANUARY 1961 NUMBER 1

EDITORIAL

THE New Year is traditionally a time for self-examination, conducted—according to temperament—in a spirit of humility or of satisfaction. The commencement of our twelfth volume, the first to be opened since the amalgamation of The Institute of Physics and The Physical Society, provides an opportunity to consider the purpose of this journal and to attempt a statement of some principles to guide its future progress.

To begin, what do we mean by applied physics? The term is difficult to define exactly, if only because all physics must eventually be applied or discarded as abstractions and speculations grow first into experiments and then into technology. The beginning and end of this process do not fall within our purview. This journal is not the proper home for a theoretical paper on the strangeness of nuclear particles or on the interpretation of radiographs taken for quality control in a welding shop. In the middle of the range there are, however, many topics of concern to us. The design, performance and data-processing system of a bubble chamber used to study the strange particles include a great deal of applied physics; discussion of the physical basis of new quality-control techniques using nuclear magnetic resonance would be equally acceptable.

In short, we are primarily a journal of physics; that is not to say that every paper submitted to the editor must contain scientific ideas of striking originality or experimental methods of absolute novelty. There will always be some contributions of this eminent standard and it is important that they should be given a permanent home in readily accessible form. Anyone who writes one or two such papers in a lifetime will make his mark, but many members of the Institute and Society can make a valuable contribution to the literature and to the work of their fellow physicists by committing to print their researches on topics which, if not at the very boundaries of knowledge, have not yet been absorbed into the routine practice of an industry or a science. We are, in fact, a Journal of Applicable Physics.

Of all the people whose efforts give life to a scientific journal, the authors and the editor are the most important since they determine its character and its status—drawing heavily on the support of readers and advertisers! The author has something to say and the editor's duty is, with certain limitations, to let him say it. Eccentricities of literary style or scientific outlook are not in themselves grounds for rejecting a paper, though the speed of publication will be greatly increased by reasonable conformity with current practice in these regards. The author should, for example, see that his text is accurate in spelling, correctly punctuated and free from obscurities or ambiguities. He should also study the journal's recommendations as to abbreviations, symbols and bibliographical references. It is unfortunate that no two societies agree exactly on the matters; consequently there is no uniformity among authors and most papers need some editorial attention in preparation for delivery to the printer.

Before that stage is reached, however, the advice of a referee will have been sought. In judging the scientific merit of the paper, he will verify the mathematical reasoning and the deductions from experimental data. He will also suggest whether expansion or partial re-writing is necessary and whether any of the text can be curtailed without loss of clarity. Some papers are rejected, but only after anxious consideration of referees' reports. Every student knows how Waterston's famous paper on the kinetic theory of gases lay unpublished for many years because the referees of 1845 did not recognize its merits. Every editor knows that story too, and takes good care to avoid a repetition.

All that has been written here would apply equally to a great number of scientific publications; what are the distinctive features of our own journal? First, its range. At a time when specialized periodicals devoted to particular topics or techniques are being born every week, we take the whole of applied (or applicable) physics for our scope. Specially commissioned reviews, dealing with recent research in a wide variety of subjects, will in future appear every month; sometimes they will be reports of conferences and sometimes the work of individual experts writing on their own specialities.

Speed of publication is another important characteristic of this journal. The centralization of editorial work formerly done in different places by the Institute and the Society allows greater resources to be deployed, particularly when a large task for one or other of the journals has to be completed quickly. Referees have agreed to give prompt consideration to papers sent for assessment and—notwithstanding their many other commitments—generally manage to return manuscripts with their comments in only a few days. Authors are being asked to help in reducing the length of the proof-reading stage. The effect of these measures is that the *British Journal of Applied Physics* can now offer publication at least as quickly as any other journal of comparable scientific standard.

Book reviews—an essential aid in judging the voluminous literature of applied physics—are already a distinctive feature of this journal and will be offered in greater numbers during the coming year. The preparation of book reviews is one of the hardest of literary exercises, but many members of the Institute and Society regularly devote abundant labour to the prompt and expert criticism of publishers' offerings. Several review supplements (which have in the past been welcomed by both readers and publishers) will be included in our present volume.

It is not enough for a periodical such as this merely to maintain its scientific and literary standards. Our subject is advancing rapidly into new and exciting realms; the journal must follow—indeed, it must sometimes go in front, marking the road for fresh advances. The money and effort required for this programme are substantial but, reinforced by the encouragement and active support of the membership, they will yield great rewards in the advancement of learning and the elevation of the profession of physicist.

J. M. A. Lenihan,
Chairman, British Journal of Applied Physics Sub-Committee.

Summarized proceedings of a conference on stress analysis —University College of North Staffordshire, April 1960

Abstract

The Annual Conference of the Stress Analysis Group of The Institute of Physics was held at the University College of North Staffordshire, Keele, Staffordshire, from 11th to 13th April 1960. The papers, which were concerned primarily with polymers and fibres, are summarized in this article.

Rubber and rubber-like polymers

THE first two papers were concerned with photo-elasticity in rubber or rubber-like polymers. Dr. D. W. SAUNDERS (British Rayon Research Association, Manchester) presented a review paper which dealt with a theoretical molecular picture of what happens during deformation of a polymer in the glassy state and in the rubber-like state. In the glassy state it is usually thought that deformation involves distortion of covalent bond angles and lengths and the photo-elastic effects are probably associated with these distortions. In the rubber-like state, however, there is a well-defined molecular model—the kinetic theory of rubber-like elasticity—which gives a good quantitative description of the stress–strain and photo-elastic properties in terms of simple molecular parameters. This theory, which was expanded, is essentially a finite strain theory at constant volume. The main conclusions are that differences between any two principal stresses or any two principal refractive indices, in a sample subjected to a pure homogeneous strain, are proportional to the difference of the square of the corresponding principal extension ratios. Thus whilst Hooke's law is not, in general, obeyed, Brewster's law of photo-elasticity is, even for finite strains. The theoretical equations include parameters describing the degree of cross-linking, the flexibility of the individual polymer chains and the optical properties of a single chemical unit of the polymer together with observable quantities such as the mean refractive index and temperature. The theory was shown to be in good agreement with the measured properties of natural rubber and gutta percha vulcanites in the rubber-like state. Extensions of the theory to polymer solutions were also discussed.

Mr. E. H. ANDREWS (Natural Rubber Producers Research Association, Welwyn, Herts.) described finite-strain experiments in which the stresses around a crack in a thin sheet of rubber were measured photo-elastically using a technique previously described (Andrews 1957) and a high powered microscope. Two alternative procedures were adopted, both of which utilized a compensator cut from the same sheet of rubber as the test piece. In one procedure the separate stresses were obtained by integration of the principal stress map and in the other the principal extension ratios were deduced from measurements of the change in thickness of the specimen. The results assume a simple relation

$$\frac{1}{W} = \frac{9}{T}(x + \alpha)$$

between W, the stored energy, and x the distance from the crack tip along the axis of symmetry, where T is a parameter describing the state of strain in the bulk of the test piece and α is a constant related to the unstrained radius of the crack. This equation describes the results for all experiments and is of the form predicted by infinitesimal-strain theory for small values of x.

Glass fibres

The next group of papers was concerned with the preparation and utilization of high strength glass fibres and rods. Dr. W. F. THOMAS (University College of South Wales and Monmouthshire, Cardiff) described experiments in which the strengths of fibres and rods of diameters from 2×10^{-4} in. to 10^{-1} in. measured immediately after drawing were shown to be up to 530 000 lb/in^2 for all sizes. The variability was, however, greatest for the thick rods, the majority of measurements being within the range 200 000–530 000 lb/in^2. This variability could be reduced by holding the molten glass at a high temperature before drawing, when most of the bubbles could be removed from the surface, and by selection of test samples containing very few bubbles. These specimens always had a high strength.

Dr. D. G. HOLLOWAY (University College of North Staffordshire) gave a review of some of the simple experiments which have established the importance of surface contamination in the production of defects on glass fibres (Holloway 1959). The mechanisms whereby these defects, local inhomogeneities in composition and structure, lead to high stresses in their neighbourhood were discussed. It was suggested that it is the local concentration of stresses, at sharp corners on an arbitrarily shaped defect, arising through differences in the thermal expansion or the elastic coefficients, which results in sufficiently high local stresses to nucleate fracture. Simple qualitative evidence in support of this hypothesis was cited and photomicrographs of cracks appearing under thermal stress and under applied stress were shown.

Mr. P. E. JELLYMAN and Mr. D. A. RICHARDSON (Pilkington Bros. Ltd., St. Helens) spoke about some of the problems of utilizing the strength of glass fibres commercially, which have an average strength of about 200 000 lb/in^2, compared with the undamaged strengths in excess of 500 000 lb/in^2 found by the previous speakers, the difference being attributed to surface damage during manufacture. Because of the importance of this surface damage the individual filaments are surrounded by a protective resilient layer such as silicone rubber. The paper was concerned primarily with the mechanism of stress transference between resin and glass, with dynamic fatigue and creep in laminates and with their deterioration in strength in humid conditions. Possible ways of increasing the stiffness of laminates, by techniques such as increasing the Young's modulus of the glass, were also discussed.

Textile fibres

Dr. P. T. SPEAKMAN (Leeds University) described experiments to show how the Young's modulus of a polymer (wool keratin) depends on the degree of cross linking, which could be varied by alkali treatment. The Young's moduli were measured directly using an extensometer or by a

dynamic method in which a weighted fibre was vibrated in resonance. The moduli were fitted to Flory's equation (Flory 1956) for anisotropically cross linked polymers.

In the next paper Mr. R. J. E. CUMBERBIRCH (British Cotton Industry Research Association, Manchester) described experiments on the effect of molecular chain lengths on the mechanical properties of monofilaments and followed this by a theoretical interpretation of the results. Monofilaments with a wide range of molecular chain lengths, prepared from cellulose acetate, were stretched varying amounts under boiling water to give a wide range of molecular orientations, before conversion to cellulose filaments with sodium hydroxide. For filaments stretched a given amount it was shown that both specific stress at rupture (tenacity) and molecular orientation increase with molecular chain length, but that the tenacities of filaments of constant molecular orientation increase with chain length up to a certain value and then remain constant. For filaments containing distribution in chain length, it was found that the tenacity $\bar{T}$ of a blend of fibres is given by

$$\bar{T} = \sum_i w_i T_i$$

where w_i is the weight fraction of the component with chain length i and tenacity T_i, the close agreement between experiment and calculation being shown for two molecular orientations (birefringence) in the Table. In his theoretical treatment the author proposed that the structure of regenerated cellulose in monofilaments may be represented by basic structural units, each of which consists of pairs of adjacent

Tenacities of blends of two fractions of different molecular chain length

(53% of chain length (DP) 1164 + 47% of chain length (DP) 178)

	Dry tenacity, g/denier		*Wet tenacity, g/denier*	
	$\Delta_n = 0{\cdot}015$	$\Delta_n = 0{\cdot}025$	$\Delta_n = 0{\cdot}015$	$\Delta_n = 0{\cdot}025$
Measured	1·34	1·94	0·76	1·10
Calculated	1·35	2·01	0·69	1·09

Δ_n = birefringence.

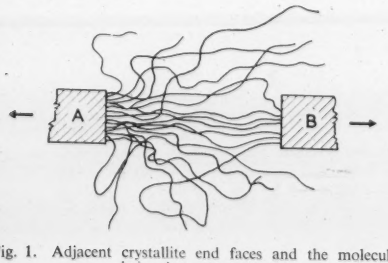


Fig. 1. Adjacent crystallite end faces and the molecular chains that connect them.

crystalline regions (crystallites) linked by chain molecules (Fig. 1). The distribution in length of these chain molecules was considered statistically, and hence the distribution of stress on the chains as the crystallites were pulled apart was calculated, the shorter chains breaking first and shedding their load to other chains. Theoretical determinations of tenacity from these statistical concepts were shown to agree closely with observation.

The difficulties in using tenacity values to predict the strength of twisted yarns were dealt with by Dr. J. W. S. HEARLE and Mr. V. M. THAKUR (Manchester College of Science and Technology). The authors suggested that the filaments at the centre of a twisted yarn are likely to be the first to break and described some experiments to test this hypothesis. They showed, however, by including fibres of various colours in the yarn, that the position of a single fibre varied from the inside to the outside at different points along the length. Breakage of the yarn could be arrested when about one half of the fibres had been broken. Five possible mechanisms that would allow this to happen were considered and it was concluded that the most probable mode of failure was one in which breakage started in the middle of one half of the yarn, spreading out to the rest of this half.

Mr. I. H. HALL (British Rayon Research Association, Manchester) described experiments to determine stress–strain curves for fibres at high rates of strain which were performed on equipment developed at the British Rayon Research Association (Lewis 1957, Holden 1959). Drawn isotactic polypropylene fibre was used which undergoes a glass-rubber transition at about 0°c and so might be expected to show pronounced changes in behaviour with increasing strain rate. The experiment was also designed to reveal any systematic errors that might occur either because the high extension velocity caused non-uniform stress distribution along the specimen, or because slippage occurred at the points where the fibre was clamped. A strain rate can be reproduced by different combinations of extension velocity and gauge length, and the different stress–strain curves should coincide unless prevented by these errors. This was done for several rates between 30 and 500 s^{-1} and it was found that provided extension velocities below 1500 cms were used and the gauge length was increased by 7 mm to correct for slippage, coincidence was achieved.

For rates from $3{\cdot}3 \times 10^{-4}$ to $3{\cdot}3 \times 10^{-2}\,s^{-1}$ and at temperatures from 90–310°K for a fixed rate of $6{\cdot}5 \times 10^{-4}s^{-1}$ stress–strain curves showed that:

(*a*) The energy to rupture and extension at break were halved as the strain rate was increased from 5×10^{-3} to $3 \times 10^{-2}\,s^{-1}$ or the temperature reduced from 280 to 270°K. This was because the material ceased to draw at the lower temperatures or higher rates.

(*b*) The initial modulus increased about 1·5 times as the strain rate was increased from 30–300 s^{-1} or the temperature was reduced from 290–250°K. Dynamic measurements at kilocycle frequencies (Sauer *et al.* 1958) show a similar change in modulus in the temperature range 310–270°K which is associated with the glass-rubber transition.

(*c*) At 90°K the extension at break was 13% and so, although the material was well below the glass–rubber transition temperature, it was not glass-like.

The data at different strain rates have been examined from the viewpoint of visco-elastic theory. Over the whole range of strain rates, for strains less than 5% the stress σ, strain ϵ, and time τ are related by the equation

$$\sigma/\epsilon = \phi(\epsilon) + \psi(\tau).$$

This is equivalent to a generalized Maxwell model containing one non-linear spring without an associated dashpot.

Mr. D. H. PAGE and Mr. P. A. TYDEMAN (British Paper and Board Industry Research Association, Kenley, Surrey) gave a brief description of the nature and manufacture of paper and outlined its rheological properties.

It should be possible to relate the mechanical behaviour of the paper sheet to three factors: the properties of the constituent fibres, the properties of the fibre-to-fibre bonds and the geometrical structure of the sheet. This relationship will obviously be of a statistical nature due to the random nature of the geometry and the distribution of the properties of the components. Initial results of measurements of the load–elongation curves of individual papermaking fibres were given and features affecting the shape of this curve were illustrated by a ciné film showing the stretching of a typical softwood fibre. Preliminary investigations into the strength of individual fibre-to-fibre bonds were described. Resulting from the development of a new technique utilizing the light microscope, it is now possible for the first time to see and measure the size and shape of individual fibre-to-fibre bonds, and this was illustrated by micrographs. This will enable factors such as the size and shape of fibre-to-fibre bonds, in isolation, to be related to their mechanical properties. Work on the structure of paper has been advanced by the application of this technique to the observation of bonds in sheets. The geometrical arrangement of bond sites has been examined and related to some papermaking variables. In addition, the technique has permitted, for the first time, the direct observation of bonds while the sheet is under stress, and the breakage of fibre-to-fibre bonds has been observed prior to rupture of the sheet. This was illustrated by a second ciné film.

Mr. F. L. WARBURTON (Wool Industries Research Association, Leeds) described experiments to determine the life under load of nylon fibres at different temperatures and moisture regains. For each experiment 50 fibres were suspended under load in a humidity cabinet controlled to $\pm 0{\cdot}2$°C and $\pm 2\%$ relative humidity. The distribution of breaking time for typical experiments is shown in Fig. 2.

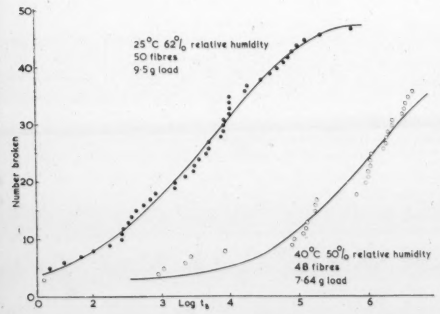


Fig. 2. Distribution of breaking times for fibres subjected to constant load.

For each set of ambient conditions experiments were made using different suspended loads, and the median breaking times or half-lives used as a preliminary measure of the effect of load on the life of the fibres. The relation between life and load so found was used to analyse the distribution of breaking times in the individual experiments in terms of a statistical distribution of minimum effective (load bearing) areas of cross section between the different fibres. It was found that the wide variation in breaking times could be accounted for on the basis of a 10% variance in cross-section area. The adequacy of this hypothesis was further examined by determining, with aid of a table of random numbers, the actual distribution of fibre cross sections to be expected when 40 fibres are drawn at random from an infinite population having this variance, several selections being made. These areas were then used in conjunction with the relation between load and half-life to construct theoretical breaking time curves, different loads being used for each group of fibres. The resultant curves are shown in Fig. 3. It will be seen that these reproduce all the features of experimental curves,

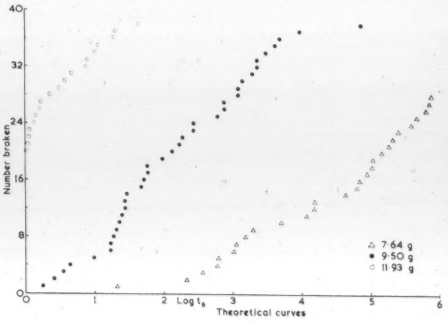


Fig. 3. Theoretical breaking time curves, based on random selection from population with variance similar to that observed experimentally.

including the step-like discontinuities thus showing that the latter can be attributed to the bunching which occurs when a finite number of fibres are drawn from an infinite population showing a continuous variation.

The data of Figs 2 and 3 and of other experimental curves have been converted to a relation between maximum stress (along the length of fibre) and breaking load assuming a Gaussian distribution of areas. It was found that the points for experiments at different loads fitted a single curve just as well for the experimental as for the theoretical data thus showing that the variance calculated from the preliminary curve between load and half-life was correct to within the limits of accuracy attainable. Master curves for the data at any one temperature have been constructed by applying correction factors to both the load and time axes. For 25°C the factor applied to the time axis was equal to the change in internal friction as calculated from experiments on the dynamic modulus at low frequencies; only a small factor was then required for the load axis to produce coincidence. These results, in conjunction with the experiments on the dynamic modulus, suggest that the effects of absorbed moisture on the mechanical properties of nylon are largely the result of a big reduction in the internal fluid friction, but much smaller changes in elastic properties also play a part.

Hot stretching is an important commercial process for treating nylon cords and experiments to determine the effect on the properties of nylon cords of time of heating, temperature and tension applied were described by Mr. W. S. SYMES (British Nylon Spinners Ltd., Pontypool). When

nylon cord is heated it tends to shrink, but under a load it creeps faster at high temperatures and these two conflicting factors govern the cord length changes. Further the modulus of the cord increases with the hot stretching tension, temperature and time. A high modulus cord has a low shrinkage when the cord is subjected to subsequent heating, such as might occur during tyre manufacture, but unfortunately thermal degradation of the nylon occurs if the heating is too severe. In practice a compromise has to be made between low shrinkage under subsequent heating, high modulus and cord strength. A fourth factor of commercial importance is that a long period of stretching reduces the rate of production of cord.

Instruments

The next two papers were concerned with the design and development of special instruments for fibre and polymer testing. Mr. D. F. DAVEY (British Nylon Spinners Ltd., Pontypool) described an instrument to measure the force required to unwind yarn from its package during textile processes. The instrument was capable of measuring tensions of from 1 mg to 5 g with a linear frequency response from 0 to 35 c/s. This was followed by a paper by Mr. A. E. EAGLES and Mr. A. R. PAYNE (Rubber and Plastics Research Association of Great Britain) which dealt with methods of converting physical stresses and strains into electrical signals for measurement and recording.

The instruments described depend on a common principle by which force, first converted to a very small movement, is detected and measured by a differential inductance or differential transformer transducer used in association with a proof ring (Van Dong and Barraud 1959).

The actual instruments developed at the Rubber and Plastics Research Association of Great Britain in which the transducer has been used were:

(1) an electronic tensile tester for measuring stress–strain characteristics of polymeric material;
(2) a dynamic tester for measuring oscillatory properties of polymeric material over a wide frequency, temperature and amplitude range;
(3) a friction measuring equipment for determining the coefficient of friction between polymer and other surfaces;
(4) a torsion pendulum dynamic modulus instrument for characterizing the temperature response of polymers;
(5) a rope friction machine for measuring the friction and abrasive stresses between rope and polymer coated fabrics; and
(6) an intermittent stress relaxation equipment for investigating the oxidative breakdown of polymers at elevated temperatures.

A practical demonstration of the system in use was given during the lecture.

Miscellaneous

On the last day of the conference several papers were presented which had no connection with polymers or fibres. In the first of these Prof. J. P. DUNCAN (Department of Mechanical Engineering, University of Sheffield) described interferometric techniques that could be used in the study of the flexure of multiply connected plates. Optically flat metal plates could be prepared rapidly to an accuracy of two interference bands over an eight-inch span. The design, construction and operation of three types of Fizeau interferometer for observing deflections and surface irregularities were outlined. The interferograms enabled the deflection w of the plates to be determined sufficiently accurately for problems of statical indeterminancy, but stresses could not be calculated from the double differentiation of w determined from the interferograms and a technique for determining these slope contours for plates in flexure was described by Mr. C. J. E. BROWN (Department of Mechanical Engineering, University of Sheffield) who first reviewed existing techniques for slope measurement. Finally, he described a technique based on the work of Salet (1939) and Ikeda (1959) in which a target mounted at one principal focus of an optical system is viewed after reflection from the plate under test at a pin hole at a second principal focus of the system (Fig. 4). Observations are made for the unloaded and loaded plate. Some

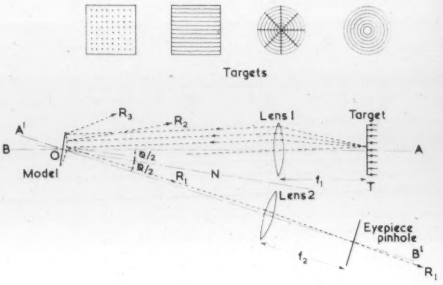


Fig. 4. Salet–Ikeda optical system for measurement of slope contours for plates in flexure.

targets of different pattern which can be used to facilitate measurement of slope contours are included in the figure. The simplest targets are the circles which give contours of equal slope and are useful for cases of circular symmetry and the straight lines which give contours of equal slope in a direction perpendicular to the orientation of the lines on the target. The apparatus which was described in detail enables slope change intervals of 0·000025 radian to be recorded without difficulty, the slope change interval being the slope difference from one equi-slope line to the next.

Mr. C. L. EMERY (Queen's University, Ontario, and University of Sheffield) next demonstrated the use of photo-elastic coatings in the study of strains in rocks and granular materials. Two basic techniques had been used: in the first photo-elastic sheet was stuck on to a plane surface of the rock under investigation and in the other, the plastic was polymerized directly on to the test surface. With rocks it was normally necessary to have a reflective surface between the plastic and the surface and this was provided either on the plastic sheet or as a reflective cement. Interpretation of observations on granular materials was discussed and in particular the use of the technique was shown for studying creep and delayed elasticity in rocks after extraction. Some results were also shown for strain patterns in large-grained aluminium test pieces. It was claimed that the technique had wide application in the field of stress analysis.

In the final paper Mr. A. DINSDALE (British Ceramic Research Association, Stoke-on-Trent) spoke about the work at B.C.R.A. on the impact strength of ceramic materials. A pendulum device was described in which a ceramic rod was

released just before contact was made with a steel anvil at the bottom of the swing. The rod was struck transversely at the mid-point so that it tended to fracture in tension. The results suggest a correlation between 'impact strength', Young's modulus and the modulus of rupture. The relation between the impact energy required to fracture a cylindrical rod and the ratio of length–diameter has been investigated. Attempts to account for the observations in terms of the various corrections to the long beam flexure formulae, required to take account of shear effects and indentation stresses, were discussed.

The conference ended with a visit to the British Ceramic Research Association, where the experiments described by Mr. Dinsdale were demonstrated. Other work connected with brittle fracture of ceramics, studies on the rheology of clays, dynamic measurement of the elastic properties of ceramics, measurement of the strength of bricks and the manufacture of high strength ceramics were also demonstrated.

Mining Research Establishment,
Worton Hall,
Isleworth, Middlesex.

C. D. POMEROY

References

ANDREWS, E. H., 1957, *J. Sci. Instrum.*, **34**, 115.

FLORY, P. G., 1956, *J. Amer. Chem. Soc.*, **78**, 5222.

HOLDEN, G., 1959, *J. Text. Inst.*, **50**, T.41.

HOLLOWAY, D. G., 1959, *Phil. Mag.*, **4**, 1101.

IKEDA, K., 1959, *Proc. 1st Japan National Congress for Appl. Mech.*, **1**, 219.

LEWIS, G. M., 1957, *Proc. Conf. on Properties of Materials at High Rates of Strain* (London: Institution of Mechanical Engineers). p. 190.

SALET, G., 1939, *Bull. Assn Tech. Maritime and Aeronautique*, **43**, 107.

SAUER, J. A., WALL, R. A., FUSCHILLO, N., and WOODWARD, A. E., 1958, *J. Appl. Phys.*, **29**, 1385.

VAN DONG, N., and BARRAUD, A., 1959, *J. Electronics and Control*, **6**, 275.

Improvements of photoelastic technique for strain measurement on flat surfaces

by G. CLYNE, A.R.C.S.T., G.I.Mech.E., H. FESSLER, M.Sc., Ph.D., A.M.I.Mech.E., and R. W. WILSON., G.I.Mech.E., University of Nottingham

Paper received 16th September 1960

Abstract

The use of the impact glue, Eastman 910, *produced stronger bonds than could be achieved with epoxy resin adhesives. Epoxy resin layers bonded on to flat polished Dural surfaces were torn by fatigue cracks in the metal without extensive bond failure. The bond strength was improved by controlled etching and hot soaking of the Dural surface (as shown by Fessler and Haines in* 1958) *before bonding.*

Impact glue bonds between epoxy resin layers and flat polished alloy steel surfaces withstood static strains of 0·3% *(slightly beyond yielding) in pure tension and* 2·5% *in eccentric tension.*

Introduction

A TECHNIQUE had previously been developed (Fessler and Haines 1958) for the measurement of static strains in aluminium alloys (Fessler and Haines 1959). The technique employed an epoxy resin to bond layers of photoelastic material on etched surfaces. It proved unsatisfactory for aluminium components subjected to alternating forces and for steel components under static forces. An impact glue (Eastman 910, by Kodak Ltd.) proved satisfactory for both these problems.

Initial bonding procedure

The manufacturers of the impact glue state that "smooth surfaces, which provide more intimate contact, bond more readily than rough poorly-fitting surfaces". Preliminary tests confirmed this and the following procedure was evolved.

The ground or bright-rolled surface was rubbed with emery papers of successively finer grades to no. 400 grade, polished with 6 and 1 μ diamond polishing paste and degreased with petroleum ether. Surface finishes of centre line average height of 5 to 6 μ in. were obtained in this manner on Dural specimens and 1 μ in. on steel specimens.

Thin pieces of epoxy resin (Araldite casting resin B) were flattened and dried by heating above the critical temperature for frozen stress and then degreased. Drops of adhesive were squeezed out of the plastic container and placed on the metal surface at approximately $\frac{1}{2}$ in. spacings. The layer was placed on these drops and the adhesive spread over the bonding surfaces by sliding the layer over the metal without the application of pressure. Finger pressure was then applied at the centre of the bonding area and gradually spread outwards to squeeze surplus adhesive to the edges of the layer.

After less than one minute the pressure was removed and the bond was allowed to cure for twenty-four hours at room temperature in a dessicator, to impede the formation of time edge stresses in the layer. The layer was carefully filed to the profile of the metal component by exerting thrust towards the metal only.

Bonds made in this manner withstood varying tensile strains in a fatigue machine and were strong enough to cause the photoelastic layer to crack in the same positions as the fatigue cracks in the Dural plates to which they were bonded. However, local bond failure occurred at the tip of the cracks and further tests were carried out to try to increase the strength of the bonds.

Dural shear tests

The results of shear tests carried out on specimens shown in Fig. 1 are presented in Table 1. The etching of specimen

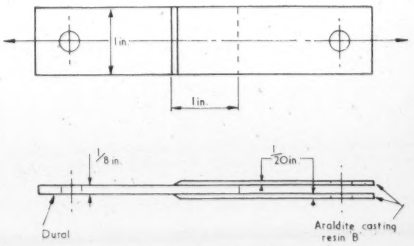


Fig. 1. Shear test specimen.

number 1.3 consisted of immersion in a special solution of 16% sulphuric acid and 2% sodium fluoride for ten minutes.

Table 1

Specimen no.	*Metal preparation*	*Spreading of adhesive*	*Type of failure*	*Failure stress (lb/in²)*
1.1	Polished	Slide	Bond	140
1.2	Polished	Scrape	Bond or end	300
1.3	Etched	Slide	End	380

The term 'scrape' refers to a new method of applying the adhesive to the metal surface. This consists of placing a pool of adhesive on to the metal surface and spreading it over the bonding area with the edge of stiff paper. This was done in an attempt to equalize (over the whole surface) the time between adhesive-epoxy contact and bonding. Trials had shown that the adhesive attacked the epoxy surface fairly

rapidly, and corrosion pits at the initial positions of the drops were suspected as weak parts of the joint because the bonding pressure at these points would be reduced.

The 'failure stresses' are the load/area of joint, purely arbitrary comparative figures, which suggest that both etching of the metal surface and distribution of the adhesive by scraping may be beneficial. With specimen 1.2 it could not be ascertained whether the bond failed before one of the epoxy plates fractured.

Dural surface tests

Four plain tensile specimens of the size shown in Fig. 2 had layers bonded over the greater part of their shanks.

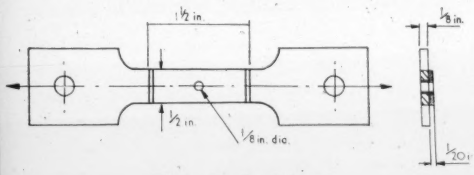


Fig. 2. Tensile specimen.

These tests are summarized in Table 2. The adhesive was spread by sliding for these four tests. Specimens 2.3 and 2.4 were etched in the same way as specimen 1.3, but specimen 2.4 was immersed in distilled water at 70° c for ten minutes after the etching solution had been rinsed off.

Table 2

Specimen no.	*Metal preparation*	*Type of failure*	*Failure stress / Yield stress*
2.1	Polished	Bond	1·26
2.2	Polished	Bond	1·44
2.3	Etched	Local at neck	1·54
2.4	Etched and soaked	Local at neck	1·31

Necking and fracture occurred near the ends of the layers. These tests confirmed that etching increased the bond strength but the controlled hot soaking, which was taken over from the technique for bonding with epoxy adhesive (Fessler and Haines 1958), had to be preceded by the correct etch. The thickness of adhesive for the polished specimens was approximately 0·0003 in., whereas it exceeded 0·002 in. for the etched ones.

Drying-out tests

The Dural specimens shown in Fig. 2 contained a sharp stress raiser. Two short fretsaw cuts located very fine cuts, produced with a serrated razor blade, diametrically opposite, on the diameter of the hole normal to the direction of loading. Sawing with the razor blade was continued till it made a visible mark in the metal and in the layer.

The photoelastic layer absorbs atmospheric moisture, which causes time edge stresses. These time edge stresses could be removed by drying out at an elevated temperature if such treatment did not reduce the strength of the bonds. Specimens 3.1 to 3.3 were polished and the adhesive was spread by sliding. The nominal failure stress is the applied load divided by the gross cross-sectional area of the necked portion. It is apparent from Table 3 that drying out at 65° c weakens the bond very severely. Time edge stresses in the layers must therefore be avoided by ensuring that the layer is stress-free before bonding, and by storing the bonded specimen in a dessicator.

Table 3

Specimen no.	*Time at 65° c*	*Type of failure*	*Nominal failure stress / Yield stress*
3.1	2 hours	Bond	Failed in filing
3.2	1 hour	Bond	0·62
3.3	None	Local	0·85 tears layer

The test on specimen 3.3 showed that the bond strength in static loading was comparable with the fatigue strength, because the layer was torn in both cases.

Layer surface tests on Dural

Examination after failure of the metal surface of some of the etched specimens revealed partial adhesive covering. This indicated that, with etched aluminium surfaces, the bond between the adhesive and the metal was of comparable strength with the bond between the adhesive and the epoxy layer. The effect of preparation of the layer surface was studied with specimens 4.1 to 4.3 of the shape shown in Fig. 2.

Table 4

Specimen no.	*Layer surface*	*Adhesive on*	*Type of failure*	*Nominal failure stress / Yield stress*
4.1	Polished	Metal	Tears layer	0·91
4.2	Emery 400	Metal	Tears layer	0·87
4.3	Emery 400	Layer	Tears layer	0·91

Table 4 shows the results of these tests. All metal surfaces were etched and soaked according to the procedure for bonding with epoxy adhesive (Fessler and Haines 1958). The impact glue was spread by scraping. Table 4 indicates that surface finish of the layer has little effect and that it is immaterial whether the adhesive is placed on the metal or on the layer. Very strong bonds were attained in every case and no bond failure could be detected. The etching of the Dural surface reduced its reflectivity, and high contrast photographic emulsions were required to produce satisfactory photographs of fringe patterns.

The epoxy layers used in all these tests were cast in horizontal open trays and the air-face of these castings was used as a 'polished' layer surface. The 'emery' surfaces were produced by rubbing down the lower surface of the sheets. It was noted that the impact glue reacted less rapidly with the air-face of the layers.

Fringe decay test

A plain tensile specimen of the size shown in Fig. 2 was prepared by etching and soaking and an epoxy layer was bonded to the shank. The specimen was then loaded in pure tension and maintained at a constant strain of 0·087% for sixty-eight hours, during which time no change of fringe order was detected. The strain was then increased to 0·24% and kept constant for a further seventy-eight hours, and again no change in fringe order was recorded. Both these strains were elastic.

Alloy steel tests

Various etching procedures were tried in unsuccessful attempts to achieve a strong joint between epoxy layers and flat surfaces of an alloy steel (0·31% carbon; 0·49% manganese; 2·1% nickel; 0·67% chronium; 0·21% molybdenum)

using epoxy adhesives. The initial bonding procedure described earlier proved successful for static loading of flat steel models, which contained different discontinuities of profile but no cracks.

Several unexpected bond failures led to a series of six tests of plain tensile specimens. These showed that no significant increase in bond strength was obtained by:

(*a*) using emery paper or etched metal surfaces instead of polished surfaces, or,
(*b*) spreading the adhesive by scraping instead of sliding.

The unexpected bond failures were later shown to have been caused by an oxide film, which formed on the polished steel surfaces while they were stored for several weeks before bonding.

When the 'initial bonding procedure' was adopted with freshly polished pure tension specimens it was found that the bond failed at 0·3% strain, which was slightly greater than the yield strain of 0·29%. In cases of eccentric tension (where the greatest bending stress was 5·1 times the mean tensile stress) the bonds withstood strains of up to 2·5%.

Two long-term tests on pure tension specimens of another alloy steel showed no change in the fringe orders when a strain of 0·077% had been maintained for 120 hours and a strain of 0·15% had been maintained for eighty hours. Both these strains were elastic.

Conclusions

The impact glue Eastman 910 was found superior to epoxy adhesives for bonding layers of epoxy resin on to aluminium and steel surfaces.

With Dural specimens the best results were obtained by controlled etching and hot soaking of the surface (Fessler and Haines 1958) prior to spreading the adhesive by scraping on either surface. Heating to 65° C destroyed the strength of the bonds.

The surface finish of alloy steel does not appear to affect the bond strength, provided that bonding takes place before a significant oxide film is formed on the surface.

Acknowledgments

The authors wish to thank Dr. H. K. M. Lloyd, Senior Lecturer in Metallurgy, for his advice on the metallurgical aspects of the problem. The work on Dural surfaces forms part of a project supported by a grant from the Department of Scientific and Industrial Research.

References

FESSLER, H., and HAINES, D. J., 1958, *Brit. J. Appl. Phys.*, **9**, 278, 282.
——, 1959, *Aero Quarterly*, **10**, 230.

Growth of crystals in glass

by J. G. MORLEY, M.Sc., A.Inst.P.,* Pilkington Bros., St. Helens, Lancs.

Paper received 16th September 1960

Abstract

The devitrification characteristics of five Na_2O–SiO_2 glasses have been examined by a new method, and some empirical relationships, applicable over the range of compositions studied, have been established. Comparisons are made with results given in other published work.

Introduction

A GLASS will show a tendency towards the formation of crystals over a particular temperature range, which varies according to its composition. Suitable heat treatment within this range will cause crystals to grow in the glass, which is then said to be devitrified. The upper limit is defined as the liquidus temperature, and above this point crystals are unstable and will dissolve in the glass. The lower limit of the devitrification range can be regarded as the point at which the rate of crystal growth becomes negligibly small. Between this point and the liquidus temperature the rate of crystal growth has a maximum value.

A glass devitrifies with varying facility according to its composition, and this is one of the main factors limiting the composition range of practical glasses. All glasses at some stage in their manufacture have to pass through their devitrification range, and the manufacturing process must be such that the time spent by the glass in this region is as short as possible. A glass composition chosen for production must have the minimum devitrification tendency consistent with its other desirable properties.

Measurements of the rate of crystal growth as a function of temperature have been made by a number of observers, including Dietzel (Zschimmer 1929), Mullensiefen and Zschimmer (1931), Preston (1940), Grauer and Hamilton (1950), Swift (1947a) and Milne (1952). Littleton (1931) noticed a general correlation between the maximum rate of crystal growth and the viscosity of the glass at the corresponding temperature in a number of Na_2O–CaO–SiO_2 glasses examined by Dietzel (Zschimmer 1929). Preston (1940) and Swift (1947b) devised empirical equations connecting the rate of crystal growth with the glass viscosity and the degree of cooling below the liquidus temperature; and Cox and Kirby (1947), using a different approach, also derived an equation connecting the rate of crystal growth with temperature.

Published work deals mainly with glasses of technical interest containing, in general, a large number of different

* Now at Rolls-Royce Aerophysics Laboratory, Littleover, Derby.

components and a corresponding variety of different crystalline devitrification products. It was felt worth while to examine the devitrification characteristics of a series of five Na_2O–SiO_2 glasses all lying in the tridymite stability field and having liquidus temperatures ranging from 900 to 1200° C. Tridymite is the crystalline form of silica stable between 870 and 1470° C. The viscosities of various glasses in this system have been studied by Lillie (1939), and by Shartsis, Spinner and Capps (1952).

Experimental work

The bulk of the published work on the rate of crystal growth has been obtained by a microscopic examination of crystalline spherulites in small samples of glass, which had been held at particular temperatures within the devitrification zone and then quenched. Dietzel (Zschimmer 1929) and Swift (1947a) also made measurements on the growth of a crystalline layer at the surface of their glass samples, at the lower temperature end of the devitrification range of the composition. It was found during the course of the study, that measurements of this type showed far less variability than measurements made on individual spherulites. The technique was therefore adapted for use in the present investigation. In their devitrification range the glasses studied had fairly high viscosities, low rates of crystal growth and very rare homogeneous nucleation. Because of these properties it was found possible to modify the method as used by Swift (1947a) and Dietzel (Zschimmer 1929), so that by placing a long sample in a suitable temperature gradient, the rate of increase in thickness of the crystalline surface layer could be measured throughout the devitrification range. It was established that the rate of growth of such a layer was uniform at a particular temperature. It was also established that, in the case of the growth of silica crystals, it had the same value as the rate of increase of the radius of a crystalline spherulite at the same temperature.

The samples used were in the form of rectangular blocks, about 10 cm long and 1 cm in cross section, and were ground to these dimensions using a medium grade of carborundum powder. Care was taken to remove all traces of carborundum powder after grinding. This treatment was found to give an ample number of nucleation centres at the surfaces of the glass. The samples were wrapped in 0·05 mm thick platinum–10% rhodium foil and placed in a small horizontal platinum-wound electric furnace, which had an almost linear temperature gradient in its working zone. The temperature distribution along the sample was determined from the readings of twelve platinum–platinum–13% rhodium thermocouples, which were embedded in a refractory tile and so arranged that their hot junctions were almost in contact with the base of the sample. After a time, dependent on the maximum rate of crystal growth in the glass, the sample was removed from the furnace and annealed, care being taken to ensure that the maximum temperature reached in the annealing programme was outside the devitrification zone of the glass on test. The foil was removed from the sample, which was then attached to a glass plate using Canada balsam. The upper surface of the block was then ground away until all crystalline material growing inwards from it had been removed. The depth of penetration of the crystalline surface layer growing inwards from the sides of the block was then measured at various positions along the length of the block, corresponding to different temperatures.

The rate of crystal growth temperature curves for the glasses studied are shown in Figs 1 to 5. The liquidus temperatures were taken from the phase diagram of the Na_2O–SiO_2 system published by Kracek (1930, 1939). The primary phase in all these glasses was identified as cristobalite (the crystalline form of silica stable between 1470 and 1713° C), by x-ray diffraction photographs, not tridymite as would have been expected from the phase diagram. This anomalous

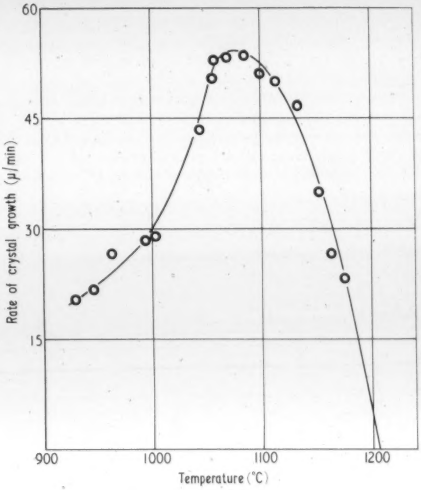


Fig. 1. Relationship between temperature and rate of crystal growth for glass of composition 82 mol% SiO_2, 18 mol% Na_2O.

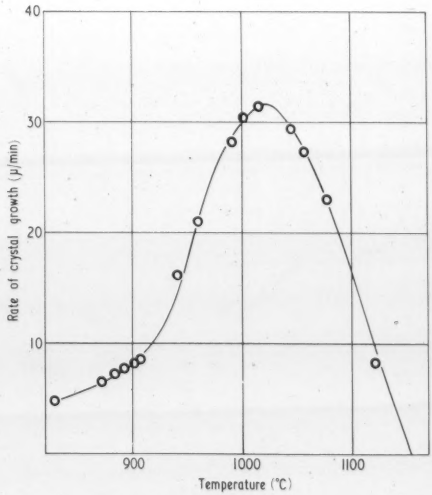


Fig. 2. Relationship between temperature and rate of crystal growth for glass of composition 81 mol % SiO_2, 19 mol % Na_2O.

behaviour is also found in Dietzel's (Zschimmer 1929) work on Na_2O–CaO–SiO_2 glasses. It was possible to measure the rate of growth of the secondary phase, sodium disilicate, in the case of glasses given in Figs 3, 4, and 5.

From data published by Lillie (1939) and by Shartsis, Spinner and Capps (1952), and from measurements made on the rate of elongation of heated uniform fibres under known loads, it was possible to derive values of viscosity at temperatures within the devitrification range of the glasses studied. Littleton (1931) pointed out that, in the case of some of the glasses examined by Dietzel (Zschimmer 1929), the maximum rate of growth of the crystal phase was approximately pro-

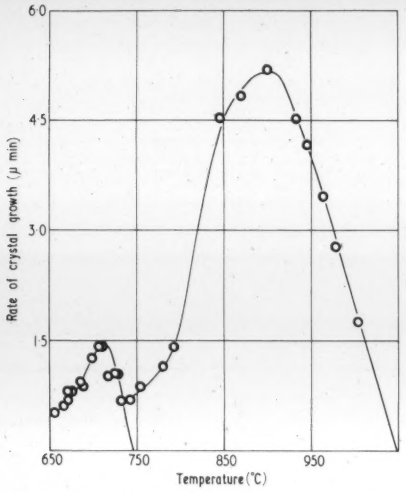


Fig. 3. Relationship between temperature and rate of crystal growth for glass of composition 79 mol% SiO_2, 21 mol% Na_2O.

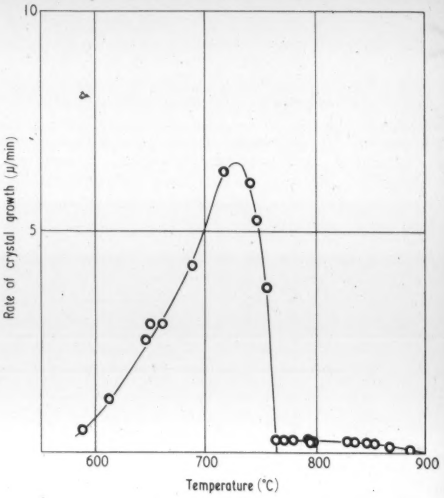


Fig. 5. Relationship between temperature and rate of crystal growth for glass of composition 76·5 mol % SiO_2, 23·5 mol % Na_2O.

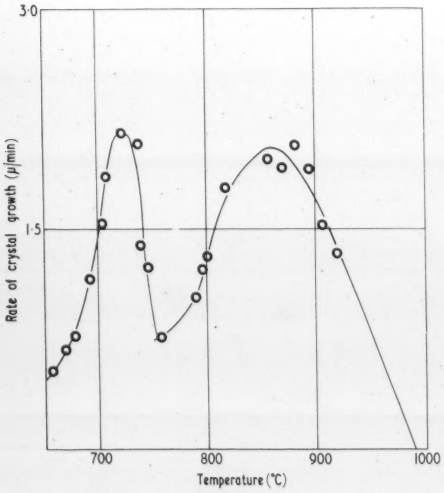


Fig. 4. Relationship between temperature and rate of crystal growth for glass of composition 78 mol% SiO_2, 22 mol % Na_2O.

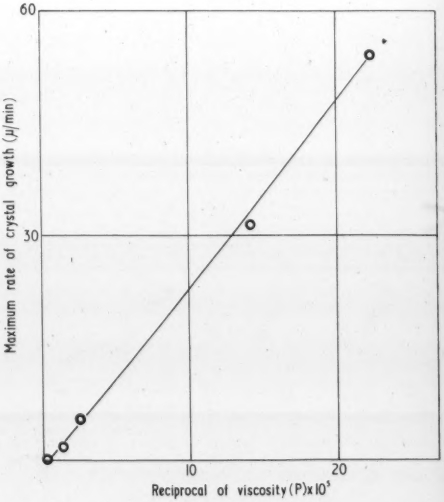


Fig. 6. Relationship between fluidity and maximum rate of crystal growth.

portional to the fluidity of the glass at that temperature. The same relationship is found to hold in the case of the glasses studied here (Fig. 6).

None of the devitrification curves obtained in the present study can be adequately described in terms of the empirical relationships established by Preston (1940) and Swift (1947b) and by Cox and Kirby (1947) for other glass compositions.

A linear relationship was found to exist between the logarithm of the maximum rate of crystal growth and the reciprocal of the liquidus temperature (Fig. 7). The phase diagram showing the change in the liquidus temperature with composition for the system Na_2O–SiO_2 has been published by Kracek (1939): therefore, if the empirical relationship shown in Fig. 7 is assumed to hold for glasses of higher silica content, the values of maximum rates of crystal growth can be predicted for Na_2O–SiO_2 glasses containing the same amounts of Na_2O as the Na_2O–CaO–SiO_2 glasses studied by Dietzel (Zschimmer 1929). A comparison can be made between Dietzel's results and those suggested by the present study, since these values can be regarded as end points on curves for Dietzel's data plotted so as to show the change in maximum rate of growth of cristobalite with CaO content in his glasses. This is illustrated in Fig. 8.

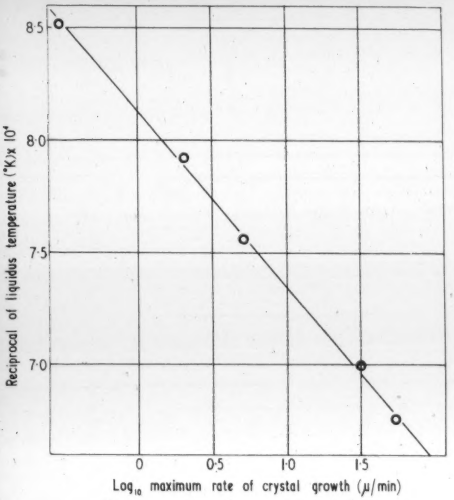


Fig. 7. Relationship between the reciprocal of the liquidus temperature and the logarithm of the maximum rate of crystal growth.

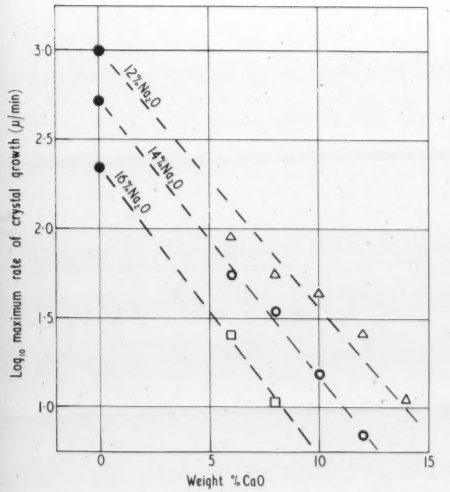


Fig. 8. Comparison between Dietzel's results and values suggested by the present study: ● = values suggested by present study, △ ○ □ = Dietzel's results.

Conclusions

The method described for measuring the growth of crystals in glasses was found to be very suitable for compositions in which homogeneous nucleation was rare, and where the viscosities in the devitrification range were sufficiently high so as to prevent flow of the glass. These properties are usually found in the case of commercial glasses, for which the method has been found suitable; and the fairly large sample used successfully eliminated the effects of small scale inhomogeneities. Empirical equations describing the variation with temperature of the rate of crystal growth, which have been devised for other glass compositions, cannot be applied to the results obtained in the investigation. It was shown that for the glasses examined, the logarithm of the maximum rate of crystal growth varied linearly with the reciprocal of the liquidus temperature. This relationship, extrapolated to glasses of higher silica content, predicts results which can be compared with published work.

Acknowledgments

The author is grateful to the Directors of Pilkington Bros. for permission to publish this paper, and also wishes to acknowledge the helpful advice of Dr. N. Davy of the University of Nottingham.

References

Cox, S. M., and Kirby, P. M., 1947, *Nature (London)* **159**, 162.
Grauer, O. H., and Hamilton, E. H., 1950, *J. Res. Nat. Bur. Stand.*, **44**, 495.
Kracek, F. C., 1930, *J. Phys. Chem.*, **34**, 1588.
——, 1939, *J. Amer. Chem. Soc.*, **61**, 2869.
Lillie, H. R., 1939, *J. Amer. Ceram. Soc.*, **22**, 367.
Littleton, J. T., 1931, *J. Soc. Glass Technol.*, **15**, 286.
Milne, A. J., 1952, *J. Soc. Glass Technol.*, **36**, 275.
Mullensiefen, W., and Zschimmer, E., 1931, *Glastechn. Ber.*, **9**, (5), 280–307.
Preston, E., 1940, *J. Soc. Glass Technol.*, **24**, 101.
——, 1940, *J. Soc. Glass Technol.*, **24**, 139.
Shartsis, L., Spinner, S., and Capps, W., 1952, *J. Amer. Ceram. Soc.*, **35**, 155.
Swift, H. R., 1947a, *J. Amer. Ceram. Soc.*, **30**, 165.
——, 1947b, *J. Amer. Ceram. Soc.*, **30**, 170.
Zschimmer, E., 1929, *J. Soc. Glass Technol.*, **13**, 76.

Method for calculating boundary value problems in heat conduction for the cylindrical cavity and the half space, by means of convolution integrals

by D. W. JORDAN, B.Sc., Mining Research Establishment, National Coal Board, Isleworth, Middlesex

Paper received 21*st July* 1960

Abstract

A convenient method for obtaining approximate solutions to certain types of boundary value problems in heat conduction is by the use of convolution integrals. The method can be used when no analytical solutions exist, or when they are too complicated to apply. An advantage is that attention can be confined to the temperature variation at one point, say, on the surface, whereas finite difference methods require the temperatures over the whole field to be computed. In this paper surface temperatures are principally considered, and a table of the function required for computations relating to a cylindrical cavity is presented.

1. Introduction

Some time ago the author found it necessary to calculate, as part of a problem in mine ventilation being solved on a digital computer, the surface temperature as a function of time on a cylindrical cavity, convection and radiation taking place simultaneously into a medium of varying temperature, with time-varying heat transfer coefficient. The problem of cooling by moisture evaporation, which involves a non-linear boundary condition, was also to be taken into account. The present method was devised and successfully used to solve problems of this type, where no analytical solution exists or where the analytical solution is too complicated to apply. It is based straightforwardly on a step-by-step application of a variant of Duhamel's theorem (Carslaw and Jaeger 1959, p. 30), as follows:

If $\theta = k(x, y, z, t)$ represents the temperature at time t in a solid, the initial temperature of which is zero, with the heat flux into the surface subsequently becoming unity, then the solution when the flux into the surface is a function of time $f(t)$, is given by:

$$\theta(x, y, z, t) = \int_0^t f(t - \lambda) \frac{\partial}{\partial \lambda} k(x, y, z, t) d\lambda \qquad (1)$$

Eqn (1) is the solution of the equation of heat conduction:

$$\partial\theta/\partial t = \kappa \nabla^2 \theta \qquad (2)$$

where κ is the diffusivity; for the boundary condition

$$K(\partial\theta/\partial n) = f(t) \qquad (3)$$

for $t > 0$, where K is the conductivity, and the derivative is taken in the direction of the outward normal to the bounding surface. The initial temperature is zero everywhere. $k(x, y, z, t)$ satisfies the same equation, with a unit function boundary condition:

$$K(\partial k/\partial n) = 1 \qquad (4)$$

for $t > 0$.

The result above follows because $f(t)$ can be continuously approximated by the sum of a sequence of step functions starting at successive intervals of time; an increment at time $t = \lambda$ giving a solution $k(x, y, z, t - \lambda)$ multiplied by the amplitude of the step, $\{(df/dt)dt\}_{t=\lambda}$. These solutions can then be added, or integrated, (owing to the linearity of the equation), and the result is equivalent to Eqn (1).

Carslaw and Jaeger (1959, p. 32) show that theorems of this type apply to a wider class of diffusion equation than Eqn (2). Results of this type are well known in other fields of physics and are not peculiar to heat conduction problems. They are merely a special type of integral representation arising from the resolution of the 'input' into the sum of step functions, rather than, for example, the sum of harmonic terms or delta functions. This representation is very convenient because the integral is over a finite range, and the singularities are relatively mild. However, although this is well known, the integrals do not seem to be used for computations, partly because the usual statement of Duhamel's theorems is in terms of some limited type of boundary condition, such as a linear radiation condition, and partly because, except in the case of heat flow in a plane half-space, the function k is difficult to tabulate. It will be shown that the expression in terms of flux at the surface enables problems involving a very wide class of boundary conditions, including non-linear radiation-type problems, to be computed.

In the following section the process of calculation is illustrated with reference to some problems of the half-space and compared with known solutions, and in a later section problems on the surface of a circular cylindrical cavity are considered, with a tabulation of the function appropriate to this case. Although possible improvements in the process as presented are almost obvious, and can be drawn as required from numerical integration techniques (Hartree 1958) and the theory of the numerical solution of integral equations (Fox and Goodwin 1953), stress is laid here on the simplest possible method, which will give quite satisfactory results for many purposes.

2. Heat flow in a half-space

If the co-ordinate x is measured perpendicularly to the surface into the body, the function k required by Eqn (1), giving the temperature arising from constant flux into the surface, is (Carslaw and Jaeger 1959, p. 75):

$$k(x, t) = \frac{2}{K}\left[\left(\frac{\kappa t}{\pi}\right)^{1/2} \exp(-x^2/4\kappa t) - \tfrac{1}{2}x \operatorname{erfc}\left\{\frac{x}{2(\kappa t)^{1/2}}\right\}\right] \qquad (5)$$

Confining the discussion to surface temperatures, we need

$$k(0, t) = k(t) = (2/K)(\kappa t/\pi)^{1/2} \tag{6}$$

In the above, K is the conductivity and κ the diffusivity of the solid. We write, conveniently,

$$4\kappa t/\pi K^2 = \tau \tag{7}$$

so that

$$k(\tau) = \tau^{1/2} \tag{8}$$

(τ has dimensions (flux)$^{-2}$), and Eqn (1) becomes

$$\theta(\tau) = \int_0^{\tau} f(t - \lambda)\frac{d(\lambda^{1/2})}{d\lambda}d\lambda \tag{9}$$

where $\theta(\tau)$ represents the surface temperature. With a view particularly to obtaining a procedure suited to the cylindrical case, we shall not evaluate the derivative in Eqn (9), but replace the integral by its simplest possible finite difference form, using only the first differences, with a basic time interval $\delta\tau$:

$$\theta_n = \sum_{s=1}^{n} \overline{f_{n-s+1}}\Delta_s(\tau^{1/2}) \tag{10}$$

where

$$\overline{f}_r = \tfrac{1}{2}(f_{r-1} + f_r) \tag{11}$$

$$\Delta_s(\tau^{1/2}) = (r\delta\tau)^{1/2} - \{(r-1)\delta\tau\}^{1/2}. \tag{12}$$

Eqn (12) is simply a table of the first differences of $\tau^{1/2}$; f_r and θ_r are the values of the flux and surface temperature at time $r\delta\tau$, at the ends of the rth interval. The formula is exact when f is constant over each interval. $d(\lambda^{1/2})/d\lambda$ is infinite at $\lambda = 0$, but this does not cause the representation (10) to break down, provided that f can reasonably be represented by its mean over the nth interval.

To test the adequacy of Eqn (10), when the flux given is as a function of time, a comparison of the exact with the approximate solution is given in Table 1, for $^1(\tau) = \tau^{1/2}$, τ and $\tau^{3/2}$. In these cases, the fractional error is independent of the size of the interval and depends only on the number of steps taken. The interval chosen here is $\delta\tau = 0{\cdot}1$.

Problems would not often occur in which the flux was given as a function of time, and this illustration is presented to exhibit the order of accuracy provided by the formula (10). For many purposes it appears that satisfactory accuracy might be expected after about five steps, provided that the flux did not vary too rapidly.

The method will now be tested in a problem where neither flux nor temperature is known in advance. We shall calculate the surface temperature of a plane half-space, initially at zero, receiving radiation from a medium at unit temperature. The exact solution (Carslaw and Jaeger 1959, p. 72), is

$$\theta(\tau) = 1 - \exp(\tfrac{1}{4}\pi h^2\tau)\operatorname{erfc}\{(\tfrac{1}{4}\pi h^2\tau)^{1/2}\} \tag{13}$$

where h is the heat transfer coefficient, defined by:

$$\text{flux into surface} = h\{1 - \theta(\tau)\}. \tag{14}$$

(This differs from the definition of h given by Carslaw and Jaeger.)

Since θ appears linearly, the Eqns (10) form a set which can be solved recursively, as follows. (This method was pointed out by Mr. E. L. Albasiny; the work had previously been done by the approximate methods described later.) Substituting for f into Eqn (11), by means of Eqn (14), we obtain

$$\overline{f}_r = h(1 - \tfrac{1}{2}\theta_{r-1} - \tfrac{1}{2}\theta_r)$$

and Eqn (10), for $n = 1$, becomes, since $\theta_0 = 0$,

$$\theta_1 = h(1 - \tfrac{1}{2}\theta_1)\Delta_1$$

or

$$\theta_1 = h\Delta_1/(1 + \tfrac{1}{2}h\Delta_1). \tag{15}$$

Eqn (10), with $n = 2$, gives

$$\theta_2 = \overline{f}_1\Delta_2 + \overline{f}_2\Delta_1$$

so that, substituting for $\overline{f}_1$ and $\overline{f}_2$ and solving for θ_2,

$$\theta_2 = h\{(\Delta_1 + \Delta_2) - \tfrac{1}{2}\theta_1(\Delta_1 + \Delta_2)\}/(1 + \tfrac{1}{2}h\Delta_1)$$

and so on. The solution for θ_n, when $\theta_1 \ldots \theta_{n-1}$ are known, is:

$$\theta_n = h\Big(\sum_{s=1}^{n}\Delta_s - \sum_{s=2}^{n}\bar{\theta}_{n-s+1}\Delta_s - \tfrac{1}{2}\theta_{n-1}\Delta_1\Big)/(1 + \tfrac{1}{2}h\Delta_1) \tag{16}$$

where

$$\bar{\theta}_r = \tfrac{1}{2}(\theta_{r-1} + \theta_r). \tag{17}$$

Also

$$\sum_{s=1}^{n}\Delta_s = (n\delta\tau)^{1/2}$$

by the definition of $\Delta_s(\tau^{1/2})$.

Table 1. *Comparison of theoretical with calculated surface temperatures of a plane when the flux is given*

	$f(\tau) = \tau^{1/2}$, $\theta(\tau) = \frac{1}{4}\pi\tau$			$f(\tau) = \tau$, $\theta(\tau) = \frac{1}{3}\tau^{3/2}$			$f(\tau) = \tau^{3/2}$, $\theta(\tau) = 3\pi\tau^2/16$		
τ	*Calculated*	*Theoretical*	*Percentage error*	*Calculated*	*Theoretical*	*Percentage error*	*Calculated*	*Theoretical*	*Percentage error*
0	0·000	0·000		0·000	0·000		0·000	0·000	
0·1	0·050	0·079	36·6	0·016	0·021	23·8	0·005	0·006	16·6
0·2	0·141	0·157	10·2	0·053	0·060	11·7	0·021	0·024	12·5
0·3	0·223	0·236	5·5	0·104	0·110	5·5	0·050	0·053	5·7
0·4	0·304	0·314	3·2	0·163	0·169	3·5	0·090	0·094	4·3
0·5	0·383	0·393	2·5	0·230	0·236	2·5	0·142	0·147	3·4
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
1·0	0·779	0·787	1·0	0·661	0·667	0·9	0·581	0·589	1·4

Table 2 compares the values of θ_n given by Eqn (16) {which are exact solutions of the difference Eqns (10) representing the integral Eqn (9)}, with the theoretical solution, Eqn (13), in the special case $h = 0{\cdot}5$, using an interval $\delta\tau = 0{\cdot}1$. It can be seen that, even with a totally unrefined procedure, the accuracy is very satisfactory after the first point.

Table 2. *Accuracy of the finite-difference method for a radiation problem, with* $h = 0{\cdot}5$

τ	0	0·1	0·2	0·3	0·4	0·5	0·6	0·7
θ *from Eqn*	(16) 0·0000	0·1465	0·1920	0·2260	0·2531	0·2757	0·2952	0·3124
θ_1 *theoretical*	0·0000	0·1408	0·1895	0·2243	0·2515	0·2742	0·2940	0·3113
Error	0·0000	0·0057	0·0025	0·0017	0·0016	0·0015	0·0012	0·0011

We may suppose that the heat transfer coefficient h varies with time, and the temperature from which radiation is taking place also varies with time. Then no analytical solution exists, but the integral equation (9) still holds, with its finite-difference approximation (10), and once again the difference equations arising from Eqn (10) can be solved recursively. If we call the temperature of the radiating medium $\phi(\tau)$, and put

$$\phi(r\delta\tau) = \phi_r,\ h(r\delta\tau) = h_r \tag{18}$$

and substitute $$f_r = h_r(\phi_r - \theta_r) \tag{19}$$

into Eqns (10) and (11), then the first equation becomes

$$\theta_1 = \bar{f}_1\Delta_1 = \tfrac{1}{2}\{f_0 + h_1(\phi_1 - \theta_1)\}\Delta_1$$

from which

$$\theta_1 = (\tfrac{1}{2}f_0 + \tfrac{1}{2}h_1\phi_1)/(1 + \tfrac{1}{2}h_1\Delta_1) \tag{20}$$

where f_0 is the flux at time $t = 0$. For $n = 2, 3 \ldots$ successively,

$$\theta_n = (\sum_{s=2}^{n} \bar{f}_{n-s+1}\Delta_s + \tfrac{1}{2}f_{n-1}\Delta_1 + \tfrac{1}{2}h_n\phi_n\Delta_1)/(1 + \tfrac{1}{2}h_n\Delta_1) \tag{21}$$

where all the f's in the equation are known.

When the relation defining the flux is not linear in the surface temperature θ, the difference equations cannot be solved explicitly. This may occur, for example, if the radiation takes place by Stefan's fourth power law, or if the extraction of heat is by evaporation of moisture. We may then attempt to obtain an approximate solution by other methods. With an eye to the cylindrical case, considered later, the methods will be kept as simple as possible.

Unfortunately, the author does not know of any non-linear problems of this type which have been solved, so in order to get a comparison with theory, the methods will be illustrated by the simple radiation problem above, with solution (13). All that is required for application to any other problem is to know the flux numerically as a function of surface temperature: for example, by a relation such as Eqn (19), or by one defining the rate of evaporation of moisture from the surface.

We again choose for the example $h = 0{\cdot}5$ and $\delta\tau = 0{\cdot}1$. Assuming that the surface temperature at the end of the first interval has not changed very much from zero, we may guess that the mean flux over this interval is approximately equal to h (equals 0·5). This provides the first entry in a column for $\bar{f}$. Then $\bar{f}_1\Delta_1 \simeq \theta_1$, the temperature at the end of the first interval. To calculate θ_2, f_2 is needed, but is not available, so we may try using f_1, the value at the end of the first interval:

$$\theta_2 \simeq \bar{f}_1\Delta_2 + f_1\Delta_1$$

where $f_1 = h(1 - \theta_1)$. We continue in this way, using $\bar{f}_1, \bar{f}_2, \ldots \bar{f}_n, f_n$ to calculate θ_{n+1}, and using the flux $f_{n+1} = h(1 - \theta_{n+1})$ to obtain $\bar{f}_{n+1}$. The working is shown in Table 3, with the theoretical values for comparison. These results are recognizable as first approximation to the solutions of the difference equations (10), and an iteration procedure may be used to improve on these values to some extent. At the end of the first interval we have an estimate of f_1, which gives an improved estimate of $\bar{f}_1$, previously

Table 3. *Calculation of surface temperatures for radiation into a half-space;* $h = 0{\cdot}5$

τ	$\bar{f}$	f	θ *estimated*	θ *theoretical*	*error*
0		0·5000	0·0000	0·0000	
	0·5000				
0·1		0·4209	0·1581	0·1408	0·0173
	0·4108				
0·2		0·4008	0·1985	0·1895	0·0090
	0·3927				
0·3		0·3846	0·2308	0·2243	0·0065
	0·3782				
0·4		0·3717	0·2567	0·2515	0·0052
	0·3662				
0·5		0·3609	0·2787	0·2742	0·0045
	0·3560				
0·6		0·3512	0·2977	0·2940	0·0037
0·7			0·3146	0·3112	0·0034

guessed to be h. This mean can now be used to obtain a better value of θ_1, and hence of f_1 and $\bar{f}_1$, and the process is repeated until two successive values differ by a small amount. The improved accuracy is exhibited in Table 4, where three iterations were used at each step, and may be compared with the results of Table 2.

Table 4. *The radiation problem of Table* 2, *using three iterations*

τ	0	0·1	0·2	0·3	0·4	.....
θ *estimated*	0·0000	0·1465	0·1922	0·2259	0·2526	.....
θ *theoretical*	0·0000	0·1408	0·1895	0·2243	0·2515	.....
Error		0·0057	0·0027	0·0016	0·0011	.....

The process converges to the true solution of the corresponding difference equation provided that $\frac{1}{2}h\Delta_1 < 1$, and the error per step is reduced by a factor of $\frac{1}{2}h\Delta_1$. If the heat transfer coefficient varies with time, as in Eqn (19) the corresponding parameter is $\frac{1}{2}h_n\Delta_1$.

(It has been pointed out by Mr. E. L. Albasiny that this iterative method is not necessarily the most efficient, and that where the flux is given analytically as a function of θ, as, for example, by Stefan's law, so that explicit differentiation is possible, Newton's method (Hartree 1958, p. 215) will give more rapid convergence to the solution.)

A further simplication will be noted here. It will not always be convenient to keep to the same interval, and if a solution settles down to a slowly varying function of time, the operator will naturally increase the interval. Moreover, the flux history in the remote past is of less importance, and as the solution advances it will be found possible to increase the interval relating to the more distant parts of the sum-

mation, Eqn (10), while retaining smaller intervals for the more recent section.

Tests of these methods, used with obvious modifications, where surface temperature is presented and it is required to find the flux or when the type of boundary condition changes at some moment, have proved equally successful.

3. Surface temperatures in a cylindrical cavity

The methods described above were originally used for problems in a cylindrical cavity, in connection with a study of mine ventilation mentioned in the Introduction. Analytical solutions to problems in the cylindrical co-ordinate system are of a different order of difficulty from plane problems. Where they exist, they are almost always of very complicated form, involving infinite integrals or summations over expressions involving Bessel functions, requiring a great amount of numerical work in tabulating the results. On the other hand, the *numerical* form of the solution is no more complicated than in the plane case and the present method can be expected to work equally well.

We wish to solve the following problem:

$$\frac{\partial^2\theta}{\partial r^2} + \frac{1}{r}\frac{\partial\theta}{\partial r} - \frac{1}{\kappa}\frac{\partial\theta}{\partial t} = 0 \qquad (22)$$

in the region

$$r \geqslant a$$

where a is the radius of the cavity; with the initial temperature zero, and the boundary condition for $t > 0$ given in terms of the flux $f(t)$ into the surface:

$$-K(\partial\theta/\partial r) = f(t) \text{ at } r = a. \qquad (23)$$

As in Section (1), we solve the problem with the boundary condition $f(t) = 1$ and construct from it the convolution integral defining $\theta(r, t)$. It is convenient to use dimensionless parameters

$$T = \kappa t/a^2,\ F(T) = 4af(t)/\pi^2 K \qquad (24)$$

and we obtain (Carslaw and Jaeger 1959, p. 338) for the surface temperature, which is of chief interest here,

$$\theta(a, t) = \theta(t) = \int_0^T F(T-\lambda)\frac{d}{d\lambda}I(\lambda)d\lambda \qquad (25)$$

where

$$I(T) = \int_0^\infty \frac{\{1-\exp(-w^2T)\}dw}{w^3\{J_1^2(w) + Y_1^2(w)\}}. \qquad (26)$$

The function $I(T)$ is the analogue of $\tau^{1/2}$ which appeared in the plane case.

For small intervals of time,

$$I(T) \simeq 2{\cdot}7842T^{1/2} - 1{\cdot}2337T + 0{\cdot}6961T^{3/2} - 0{\cdot}4626T^2 + 0{\cdot}3654T^{5/2} - 0{\cdot}3494T^3 \qquad (27)$$

the first term being identical with the function for the plane case; and for large intervals of time (Carslaw and Jaeger 1959, p. 339),

$$I(T) \simeq \tfrac{1}{8}\pi^2(\ln 4T - \gamma) \qquad (28)$$

where $\gamma = 0{\cdot}57722$, which is Euler's constant.

To use this function in calculations it was necessary to provide an adequate tabulation, and this is presented in the Appendix, together with its first differences and its first derivative.

As in the plane case, Eqns (10) and (12), we have

$$\theta_n = \sum_{s=1}^{n} \overline{F_{n-s+1}}\Delta_s\{I(T)\} \qquad (29)$$

where

$$\Delta_r\{I(T)\} = I(r\delta T) - I\{(r-1)\delta T\}, \qquad (30)$$

and δT is the time interval chosen for a particular problem. The methods available for solving problems of the cylindrical cavity are then exactly as for the plane case.

It is important to be able to choose the interval δT correctly, and we show how this may be done by comparison with the plane case in a radiation problem. The difficult part, requiring the smallest interval, is the very early stage of the problem, when temperatures vary most rapidly. During this stage the plane and cylindrical solutions will not be very different, and the physical time intervals suited to each case, when the transfer coefficient h is the same, will be closely similar. The solution, Eqn (13), to the plane problem depends only on the function (corresponding to the time variable),

$$\mu = \tfrac{1}{4}\pi h^2\tau = h^2\kappa t/K^2. \qquad (31)$$

We can, therefore, reformulate the plane radiation problem and its numerical solution as follows: Define a dimensionless flux such that

$$\phi(\mu) = 2f(t)/\pi^{1/2}h \qquad (32)$$

at corresponding values of μ and t. Then comparing Eqns (31) and (32) with Eqns (9) and (14), we have in the new variables:

$$\theta(\mu) = \int_0^\mu \phi(\mu-\nu)\frac{d}{d\nu}(\nu^{1/2})d\nu, \qquad (33)$$

$$\phi(\mu) = (2/\pi^{1/2})\{1-\theta(\mu)\}. \qquad (34)$$

Eqns (33) and (34) embrace all such radiation problems for the plane, and the solution is

$$\theta(\mu) = 1 - \exp(\mu)\operatorname{erfc}(\mu^{1/2}). \qquad (35)$$

Returning to the corresponding cylindrical problem, with time variable $T = \kappa t/a^2$, and transfer coefficient $H = 4ah/\pi^2 K$, so that

$$F(\tau) = H\{1-\theta(T)\}$$

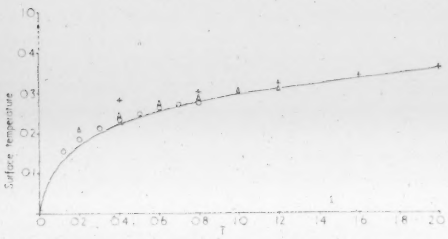
we have

$$\mu = \tfrac{1}{16}\pi^4 H^2 T. \qquad (36)$$

To determine the interval required to provide given accuracy in the cylindrical problem after some point $T = T_0$ (where T_0 is small), it is simply necessary to test what interval is required in the problem defined by Eqns (33) and (34), for the plane case, using the methods of Section 2, to provide similar accuracy at $\mu_0 = \tfrac{1}{16}\pi^4 H^2 T_0$. If this interval is $\delta\mu$, the appropriate value of $\delta T = 16\delta\mu/\pi^4 H^2$. Since the numerical method is very simple and independent of any physical constants, this test is easy to apply.

It may happen that for large values of H, the earliest interval in the table of $I(T)$, tabulated for $\delta T = 0{\cdot}1$, is too large to start the integration. In that case the expression (27) is available, which agrees with the tabulated value at $T = 0{\cdot}1$, and may therefore be used to interpolate the table for smaller values of T. The radiation problem discussed above, involving rapid temperature change at the start, is, in fact, a very stringent test of the method, and for slowly varying conditions much larger intervals can be used.

It appears also that small intervals need not be taken over the whole range. Once the steep part of the curve is surmounted, the numerical solution soon settles down to correct values. This point is illustrated by the Figure, which



Illustrating the approach of the numerical solution to the true solution for different intervals. The problem is of radiation into a cylindrical cavity with heat transfer coefficient 0·5 (unrefined process).

—— theoretical curve. ○ $\delta T = 0{\cdot}1$ △ $\delta T = 0{\cdot}2$ + $\delta T = 0{\cdot}4$

exhibits solutions of the above cylindrical radiation problem calculated for different intervals, for a value $h = 0{\cdot}5$. The rough method illustrated in Table 3 was used.

4. Conclusion

The method described for calculating surface temperatures in boundary value problems of the cylindrical cavity and the plane could also be extended to solid cylinders and spheres, and to the calculation of temperatures inside the medium, all that is necessary being the tabulation of the required functions, which are well known. Refinements in the procedure can be designed to suit particular problems, and a great improvement in accuracy can be achieved, for example, by writing in the cylindrical case, for small values of T,

$$I(T) = \tfrac{1}{2}\pi^{3/2}T^{1/2} + g(T) \qquad (37)$$

where $g(T)$ is a slowly varying function. If $F(T)$ is also found to be slowly varying over this range it may then be replaced by an interpolation polynomial. The integration over the first term in Eqn (37) is performed analytically, and over the second term numerically. However, for most problems, particularly if a computer is used, it is easier merely to take smaller intervals.

Acknowledgment

The author is indebted to Mrs. J. Webber for assistance in programming the integral $I(T)$ for an I.B.M. 650 digital computer, and for other computational work. The valuable suggestions made by Mr. E. L. Albasiny of the National Physical Laboratory, Teddington, are noted in the text. The work described here was carried out as part of the research programme of the Scientific Department of the National Coal Board and is published by permission of the Director General of Research. The views expressed are those of the author, and not necessarily those of the Board.

References

CARSLAW, H. S., and JAEGER, J. C., 1959, *Conduction of Heat in Solids*. 2nd Edn (London: Oxford University Press).

FOX, L., and GOODWIN, E. T., 1953, *Phil. Trans*, **245**, 501.

HARTREE, D. R., 1958, *Numerical Analysis* (London: Oxford University Press).

Appendix

$$\textit{Table of } I(T) = \int_0^\infty \frac{\{1 - \exp(-W^2T)\}dw}{w^3\{J_1^2(w) + Y_1^2(w)\}}$$

This integral represents $\pi^2K/4a$ times the surface temperature due to unit heat flux into the surface of a circular cylindrical cavity. It is represented asymptotically for small values of T by

$$I(T) \simeq 2{\cdot}7842T^{1/2} - 1{\cdot}2337T + 0{\cdot}6961T^{3/2} - 0{\cdot}4626T^2 + 0{\cdot}3654T^{5/2} - 0{\cdot}3494T^3.$$

The derivatives presented were obtained from the tabulation by standard formulae (Hartree 1958), and the earlier values are not thought to be reliable to the last figure.

$I(T)$ and associated quantities: $T = 0\ (0{\cdot}1)\ 2{\cdot}0$

T	$I(T)$	$\Delta I(T)$	dI/dT	T	$I(T)$	$\Delta I(T)$	dI/dT
0	0·0000						
		0·7753				0·0696	
0·1	0·7753		3·4204	1·1	2·0489		0·6720
		0·2715				0·0650	
0·2	1·0468		2·2147	1·2	2·1138		0·6289
		0·1929				0·0610	
0·3	1·2397		1·6979	1·3	2·1748		0·5916
		0·1535				0·0575	
0·4	1·3931		1·3955	1·4	2·2323		0·5586
		0·1289				0·0544	
0·5	1·5220		1·1948	1·5	2·2867		0·5294
		0·1119				0·0516	
0·6	1·6339		1·0499	1·6	2·3383		0·5034
		0·0992				0·0491	
0·7	1·7332		0·9395	1·7	2·3874		0·4799
		0·0894				0·0469	
0·8	1·8226		0·8522	1·8	2·4343		0·4588
		0·0816				0·0449	
0·9	1·9041		0·7811	1·9	2·4792		0·4394
		0·0751				0·0430	
1·0	1·9792		0·7221	2·0	2·5222		0·4218

I(T) and associated quantities: T = 0 (0·5) 20

T	$I(T)$	$\Delta\{I(T)\}$	dI/dT	T	$I(T)$	$\Delta\{I(T)\}$	dI/dT
0	0·0000						
		1·5220				0·0521	
0·5	1·5220		1·1948	10·5	4·1256		0·1020
		0·4572				0·0499	
1·0	1·9792		0·7221	11·0	4·1755		0·0978
		0·3075				0·0479	
1·5	2·2867		0·5294	11·5	4·2234		0·0939
		0·2356				0·0461	
2·0	2·5223		0·4218	12·0	4·2695		0·0903
		0·1924				0·0443	
2·5	2·7146		0·3522	12·5	4·3138		0·0870
		0·1632				0·0427	
3·0	2·8778		0·3031	13·0	4·3566		0·0840
		0·1420				0·0413	
3·5	3·0198		0·2665	13·5	4·3978		0·0811
		0·1259				0·0399	
4·0	3·1457		0·2381	14·0	4·4377		0·0784
		0·1132				0·0386	
4·5	3·2589		0·2154	14·5	4·4763		0·0760
		0·1029				0·0374	
5·0	3·3617		0·1967	15·0	4·5137		0·0736
		0·0944				0·0363	
5·5	3·4561		0·1812	15·5	4·5499		0·0714
		0·0872				0·0352	
6·0	3·5433		0·1679	16·0	4·5851		0·0694
		0·0810				0·0342	
6·5	3·6243		0·1565	16·5	4·6193		0·0674
		0·0757				0·0332	
7·0	3·7001		0·1466	17·0	4·6525		0·0656
		0·0711				0·0323	
7·5	3·7712		0·1379	17·5	4·6849		0·0638
		0·0670				0·0315	
8·0	3·8382		0·1303	18·0	4·7164		0·0622
		0·0634				0·0307	
8·5	3·9015		0·1234	18·5	4·7471		0·0606
		0·0601				0·0299	
9·0	3·9617		0·1172	19·0	4·7770		0·0591
		0·0572				0·0292	
9·5	4·0189		0·1117	19·5	4·8062		0·0577
		0·0546				0·0285	
10·0	4·0734		0·1066	20·0	4·8347		0·0563

Interferometric measurement of small angular displacements Part 2. The double-passed Jamin interferometer

by P. HARIHARAN, Ph.D., A.Inst.P., and D. SEN, M.Sc., Ph.D., National Physical Laboratory of India, New Delhi, India

Paper first received 1st *March, and in final form* 3rd *October* 1960

Abstract

When the rays emerging from a Jamin interferometer are reflected back through the instrument, fringes similar in appearance and behaviour to three-beam fringes are obtained. These fringes can be used to measure small angular displacements of one of the beam-dividing plates, with an accuracy of the order of 0·01″. *A modified set-up is also described with which angular displacements of a comparatively light, auxiliary mirror can be measured with the same degree of accuracy.*

Introduction

THE application of the principal of three-beam interference (Zernike 1950, Hariharan and Sen 1959) to obtain increased accuracy in measurements of small angular displacements has been discussed by the authors in an earlier paper (1959b), in which a three-beam interferometer modified for such measurements was described.

More recent work by the authors (1960) now shows that a system of fringes, similar in appearance and behaviour to three-beam fringes, can be obtained by reflecting the rays emerging from a conventional Jamin interferometer back through the instrument. This may be briefly termed 'double-passing' the interferometer. The application of these 'double-passed' fringes to the measurement of small angular displacements is described in the present paper.

Optical system

The optical system of the double-passed Jamin interferometer is shown schematically in the horizontal plane in Fig. 1. The two beam-dividing plates M_1, M_2 one of which (say, M_2) is the moving element, are identical, optical glass blocks, the faces of which have been worked plane and parallel. Their inner faces are coated with a $\lambda/4$ film of high refractive index giving a reflection coefficient of nearly 35%, and part of the outer faces (shown by thick lines in the diagram) is coated with an opaque, fully reflecting film of aluminium. The two beams which normally emerge on the right-hand side are reflected back through the instrument once again by the plane mirror M_3, and the double-passed fringes are viewed by means of the semi-reflecting mirror M_4 and the low-power telescope T.

Theory

The theory of the double-passed Jamin interferometer has already been derived for the case when the two beam-dividing plates M_1, M_2 are parallel in the horizontal plane (Hariharan and Sen 1960). In this section consideration is given to what happens when M_2 is rotated through a small angle $\Delta\beta$ in the horizontal plane with respect to M_1.

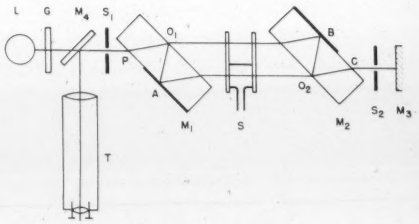


Fig. 1. Schematic diagram of the optical system of the double-passed Jamin interferometer.

L, mercury arc lamp; G, green filter; M_1, M_2, beam-dividing mirrors; M_3, plane mirror; M_4, semi-reflecting mirror; S_1, S_2, stops; S, air-cell compensator; T, viewing telescope.

From Figs 2(*a*) and (*b*), it is seen that the total optical paths of the four double-passed rays, which are formed by the division of any single incident ray, can then be expressed by the relations

$$D_{AA} = D + 2p + 2nd\cos\beta'\{\cos(\theta' + \alpha') + \cos(\theta' - \alpha')\} \quad (1)$$

$$D_{AB} = D + p + 2nd\cos(\theta' + \alpha')\{\cos\beta' + \cos(\beta' + \Delta\beta')\} \quad (2)$$

$$D_{BA} = D + p + 2nd\cos(\theta' - \alpha')\{\cos\beta' + \cos(\beta' + \Delta\beta')\} \quad (3)$$

$$D_{BB} = D + 2nd\cos(\beta' + \Delta\beta')\{\cos(\theta' - \alpha') + \cos(\theta + \alpha')\} \quad (4)$$

where d is the thickness of the interferometer plates, n is the refractive index of the material of which they are made, $\theta' = (\theta/n)$, $\alpha' = (\alpha/n)$, $\sin\beta' = (1/n)\sin\beta$, $\sin(\beta' + \Delta\beta') = (1/n)\sin(\beta + \Delta\beta)$, and p is the additional optical path introduced by a compensator located in one of the paths, as shown in Fig. 1. If one takes $\cos\Delta\beta' = 1$, and $\sin\Delta\beta' = \Delta\beta'$ to a first approximation, Eqns (1) to (4) can be rewritten in the form

$$D_{AA} = D_0 + p + 2nd\,\Delta\beta'\sin\beta'\cos\theta'\cos\alpha' \quad (5)$$

$$D_{AB} = D_0 - 4nd\cos\beta'\sin\theta'\sin\alpha' + 2nd\,\Delta\beta'\sin\beta'\sin\theta'\sin\alpha' \quad (6)$$

$$D_{BA} = D_0 + 4nd\cos\beta'\sin\theta'\sin\alpha' - 2nd\,\Delta\beta'\sin\beta'\sin\theta'\sin\alpha' \quad (7)$$

$$D_{BB} = D_0 - p - 2nd\,\Delta\beta'\sin\beta'\cos\theta'\cos\alpha' \quad (8)$$

where

$$D_0 = D + p + 4nd\cos\beta'\cos\theta'\cos\alpha' - 2nd\Delta\beta'\sin\beta'\cos\theta'\cos\alpha'.$$

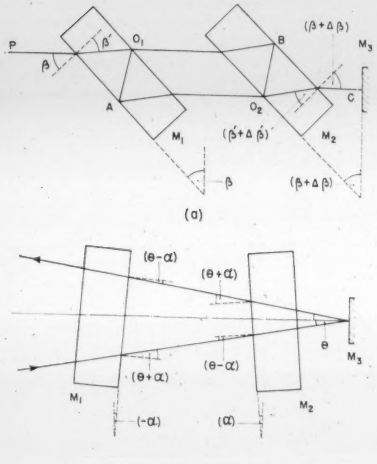


Fig. 2. Schematic diagram showing the paths of the rays in the double-passed Jamin interferometer.
(*a*) in the horizontal plane; (*b*) in the vertical plane.
(Note: the angle of incidence of the ray and the tilt of the plates M_1, M_2 in the vertical plane have been considerably exaggerated.)

The complex amplitude A_P at the corresponding point in the interference pattern is equal to the sum of the complex amplitudes of these four beams, and can be written as

$$A_P = A \exp\{-(2\pi i/\lambda) D_0\}[\exp\{-i(\phi+\delta)\} + \exp i(2\psi - \Delta\psi) + \exp\{-i(2\psi - \Delta\psi)\} + \exp i(\phi+\delta)\}] \quad (9)$$

where A is the amplitude of a single beam, $\phi = (2\pi/\lambda)p$,

$$\delta = (2\pi/\lambda)2nd\Delta\beta' \sin\beta' \cos\theta' \cos\alpha',$$
$$\psi = (2\pi/\lambda)2nd \cos\beta' \sin\theta' \sin\alpha',$$
and $$\Delta\psi = (2\pi\lambda)\, 2nd\, \Delta\beta' \sin\beta' \sin\theta' \sin\alpha',$$

so that the intensity I_P at this point, which is obtained by muliplying A_P by its complex conjugate, is given by the relation

$$I_P = A_P \cdot A_P^* = 4I_0\{\cos(\phi+\delta) + \cos(2\psi - \Delta\psi)\}^2 \quad (10)$$

where I_0 ($= A^2$) is the intensity due to a single beam.

Any slight rotation $\Delta\beta$ of M_2 in the horizontal plane results primarily in a change in the quantity δ in Eqn (10) and, therefore, in a change in the intensity distribution in the fringe system. The distance between adjacent fringes is given by the condition $(2\psi - \Delta\psi) = 2\pi$. Rotating M_2 by an angle $\Delta\beta$ in the horizontal plane also results, therefore, in a small change in the fringe spacing. However, this change is negligible under normal conditions and has no effect on the photometric setting.

As can be seen from the theoretical intensity distribution curves in Fig. 3 and the photographs in Fig. 4, alternate fringes are suppressed when $(\phi+\delta) = n\pi$. All the fringes have the same intensity only when $(\phi+\delta) = (2n+1)\pi/2$. Accordingly, this criterion can be used as a sensitive null setting to measure small changes in the angular position of M_2 in the horizontal plane.

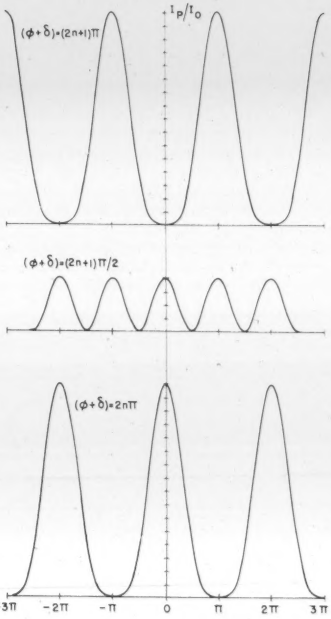


Fig. 3. Theoretical intensity distribution in the interference pattern obtained with the double-passed Jamin interferometer for various values of the phase-difference between the beams.

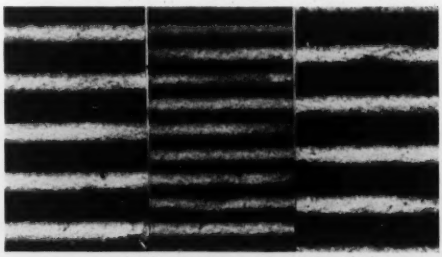


Fig. 4. Interference patterns obtained with the double-passed Jamin interferometer for various values of the phase difference between the beams
(*a*) $\phi + \delta = 2n\pi$.
(*b*) $\phi + \delta = (2n+1)\pi/2$.
(*c*) $\phi + \delta = (2n+1)\pi$.

Sensitivity

The phase difference δ, introduced between the single-passed beams by a small angular displacement $\Delta\beta$ of M_2, is given to a first approximation by the relation

$$\delta = (2\pi/\lambda)\, 2nd\, \Delta\beta' \sin\beta' \quad (11$$

which, as shown earlier (Hariharan and Sen 1959b), is a maximum when β is given by the equation

$$\sin^2 \beta = n^2 - n(n^2 - 1)^{1/2} \qquad (12)$$

that is to say, when β is about 50° for a plate of refractive index $n = 1{\cdot}5$. For a 2·5 cm thick plate, $\delta = 2\pi$ when $\Delta\beta = 5{\cdot}7''$. Since a change in $(\phi + \delta)$ of $2\pi/500$ can be detected in this interferometer (Hariharan and Sen 1960), it should be possible to detect an angular displacement of M_2 in the horizontal plane of about $0{\cdot}01''$.

As with the three-beam interferometer (Hariharan and Sen 1959b), the readings obtained with the double-passed Jamin interferometer are relatively insensitive to small angular displacements of M_2 in the vertical plane and are completely unaffected by a linear translation of M_2. In addition, the angle which M_2 makes with M_1 in the horizontal plane can be made as much as $\pm 2°$ before any lack of uniformity of the field becomes apparent. This is because the AA and BB rays in this instrument are always coincident and centred with respect to the AB and BA rays to a first approximation, when the angle which these rays make with the normal to M_3 is small. As a result, this interferometer is very easy to handle, and its adjustment is not at all critical.

a set of measurements carried out on the main slide of a precision universal measuring machine is described.

The experimental arrangement used for these measurements is shown in Fig. 5. One of the beam-dividing mirrors M_2 was clamped to the movable carriage of the measuring machine, while the other beam-dividing mirror M_1 and the plane mirror M_3, used to double-pass the interferometer, were supported by blocks of suitable height mounted on the bed of the instrument. The auxiliary semi-reflecting mirror M_4, the viewing telescope T and the compensating cell S were carried by stands mounted on two accessory bars. The air pressure in the compensating cell was varied by means of a piston P controlled by a micrometer screw and read with a mercury manometer G and a travelling microscope R.

The preliminary adjustment of the interferometer was carried out in two stages, using a white light source. In the first instance the mirror M_3 was removed, and the beam dividing plates M_1, M_2 were adjusted according to the usual procedure for a conventional Jamin interferometer, so that an eye placed behind S_2 could see a system of coloured horizontal fringes centred about the white zero-order fringe. After this, the mirror M_3 was replaced and set so that the beams emerging from the interferometer were reflected back through the instrument along the same path. This adjustment was checked by viewing the double-passed fringes through the telescope T and verifying that the central fringe in the double-passed fringe system was the zero-order white fringe. Finally, the white light source was replaced by a low-pressure mercury vapour lamp.

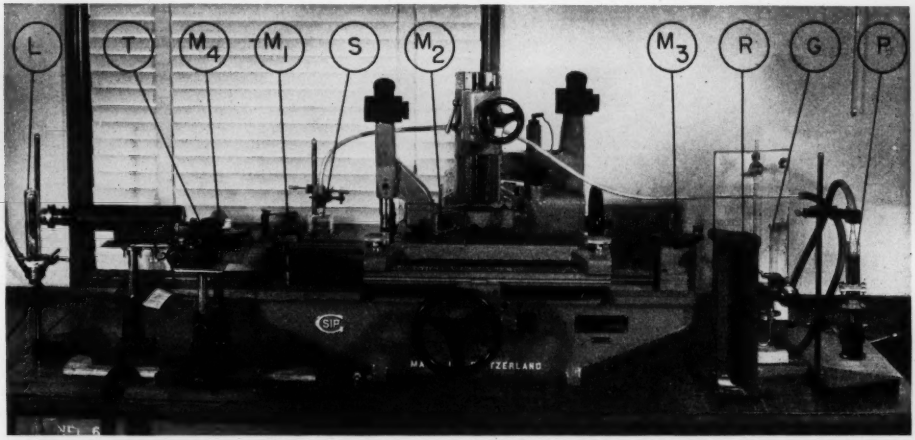


Fig. 5. Experimental arrangement used to check the straightness of the main slide of a universal measuring machine.
L, mercury arc lamp; M_1, M_2, beam-dividing mirrors; M_3, plane mirror; M_4, semi-reflecting mirror, T, viewing telescope; S, air-cell compensator; P, micrometer piston; G, mercury manometer; R, travelling microscope.

Applications

While several applications can be envisaged for the double-passed Jamin interferometer, one application for which it is particularly suitable is to measure the deviations from straightness of a precision slide. For this, one of the beam-dividing mirrors is mounted on a carriage, and the changes in the angular position of this mirror as the carriage is moved along the slide are recorded. The deviations of the slide from a straight line are then obtained by integrating these readings in the same fashion as when measurements are carried out with a reference mirror and an autocollimating telescope (Habell and Cox 1948). To illustrate the method,

The results obtained with this set-up are presented in Fig. 6, which shows the changes in the angular position of the carriage in the horizontal plane as it was moved along the slide, first in one direction, and then in the reverse direction. Two sets of measurements made on different days are presented. As can be seen, the two sets of readings agree to within $\pm 0{\cdot}04$ second of arc. The major part of this uncertainty is probably due to the variable thickness of the oil

films on the different parts of the slide. The curves show that over a range of 20 cm the carriage turns through an angle of 0·066 second of arc for every centimetre by which it is advanced. This indicates that the slide is actually an arc of a circle of radius equal to 31·3 km, corresponding to a maximum deviation from a straight line of about 0·16 μ over 20 cm. It is also apparent from these curves that the carriage of the measuring machine turns through an angle of

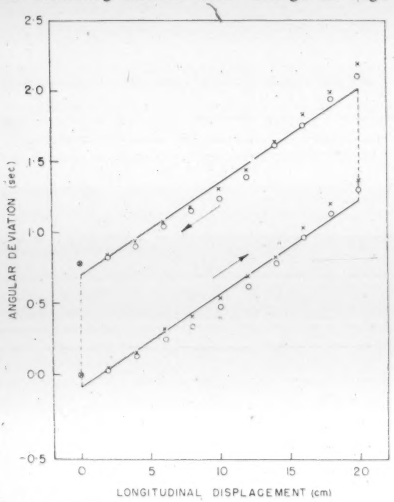


Fig. 6. Variation in the angular position of the carriage (in the horizontal plane) with longitudinal displacement along the slide.

0·8 second of arc when its direction of movement is reversed. In the case of two points where the x coordinates are separated by 20 cm, this could result in an error of as much as 0·8 μ in the y coordinates.

Modified double-passed Jamin interferometer

For problems in which it is not possible to use one of the beam-dividing mirrors as the moving element, because of its size and weight, the modified interferometer described in this section will be found quite useful. In this instrument, a single thick, plane-parallel plate M_1 is utilized in conjunction with an auxiliary plane mirror M_2, the light paths being folded as shown in Fig. 7. One face of M_1 is coated with a semi-reflecting film, while the other has an opaque reflecting strip down the middle, so that one end of it is used to divide the incident light into two beams and the other end is used to recombine them (Wolter 1955). The auxiliary mirror M_2, which is the moving element in this case, can be fairly small (about 2 cm × 1 cm) and comparatively light. The beams emerging from the interferometer are reflected back through the instrument once again by the plane mirror M_3, and the double-passed fringe system is viewed, as before, by means of the low-power telescope T.

In deriving the theory of this interferometer we assume that M_1 and M_2 are initially adjusted so that they are parallel in the horizontal plane, but that M_1 makes a small angle α with M_2 in the vertical plane, as shown in Fig. 8(*a*), and that

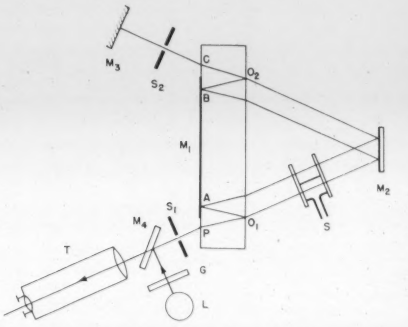


Fig. 7. Schematic diagram of the optical system of the modified, double-passed Jamin interferometer.

L, mercury arc lamp; G, green filter; M_1, beam-dividing mirror, M_2, movable plane mirror; M_3, plane mirror; M_4, semi-reflecting mirror, S_1, S_2, stops, S, air-cell compensator, T, viewing telescope.

M_3 is adjusted so that it is parallel to M_2 in the vertical plane. Any ray which is incident normally on M_3, and which is therefore reflected back along the same path, will be termed a principal ray. As before, the angle of incidence of the principal ray on M_1 in the horizontal plane is assumed to be β.

When M_2 is rotated through a small angle $\Delta\beta$ about a vertical axis, the angles of incidence of this ray on M_1 in the horizontal plane will be changed as shown in Fig. 8(*b*). The

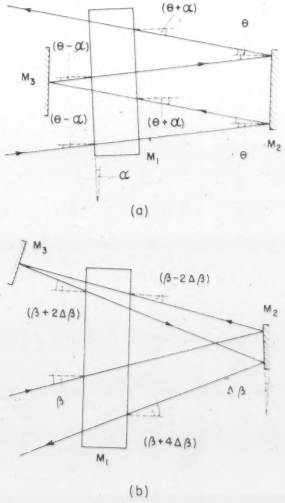


Fig. 8. Schematic diagram showing the paths of the rays in the modified double-passed Jamin interferometer.

(*a*) in the vertical plane; (*b*) in the horizontal plane.

(Note: the angle of incidence of the ray and the tilt of M_1 in the vertical plane, and the angle through which M_2 has been turned in the horizontal plane, have been greatly exaggerated for clarity.)

angle of incidence on the outward journey after reflection at M_2 will become $\beta - 2\Delta\beta$, while the angles of incidence on the return journey, before and after reflection at M_2 will become $\beta + 2\Delta\beta$, and $\beta + 4\Delta\beta$ respectively. Since $\Delta\beta$ is a very small angle, the corresponding angles of refraction within the plate can be written as $\beta' - 2\Delta\beta'$, $\beta' + 2\Delta\beta'$ and $\beta' + 4\Delta\beta'$ where $\sin\beta = n\sin\beta'$.

The optical paths of the four double-passed rays are then given by the relations

$$D_{AA} = D + 2nd\cos\beta'\cos(\theta' - \alpha') + 2nd\cos(\beta' + 4\Delta\beta')\cos(\theta' + \alpha') \quad (13)$$

$$D_{AB} = D + 2nd\cos\beta'\cos(\theta' - \alpha') + 2nd\cos(\beta' + 2\Delta\beta')\cos(\theta' - \alpha') + p \quad (14)$$

$$D_{BA} = D + 2nd\cos(\beta' - 2\Delta\beta')\cos(\theta' + \alpha') + 2nd\cos(\beta' + 4\Delta\beta')\cos(\theta' + \alpha') + p \quad (15)$$

$$D_{BB} = D + 2nd\cos(\beta' - 2\Delta\beta')\cos(\theta' + \alpha') + 2nd\cos(\beta' + 2\Delta\beta')\cos(\theta' - \alpha') + 2p \quad (16)$$

where d is the thickness of M_1, and p is the additional optical path introduced by a compensator located in the region between the plates M_1, M_2 as shown in Fig. 5.

Since $\Delta\beta'$ is a very small angle one can take $\cos 2\Delta\beta' = \cos 4\Delta\beta' = 1$, and $\sin 2\Delta\beta' = 2\Delta\beta'$, $\sin 4\Delta\beta' = 4\Delta\beta'$, so that Eqns (13) to (16) can be simplified and rewritten as

$$D_{AA} = D_0 - 4nd\,\Delta\beta'\sin\beta'\cos\theta'\cos\alpha' + 8nd\,\Delta\beta'\sin\beta'\sin\theta'\sin\alpha' - p \quad (17)$$

$$D_{AB} = D_0 + 4nd\cos\beta'\sin\theta'\sin\alpha' - 4nd\,\Delta\beta'\sin\beta'\sin\theta'\sin\alpha' \quad (18)$$

$$D_{BA} = D_0 - 4nd\cos\beta'\sin\theta'\sin\alpha' + 4nd\,\Delta\beta'\sin\beta'\sin\theta'\sin\alpha' \quad (19)$$

$$D_{BB} = D_0 + 4nd\,\Delta\beta'\sin\beta'\cos\theta'\cos\alpha' - 8nd\,\Delta\beta'\sin\beta'\sin\theta'\sin\alpha' + p \quad (20)$$

where

$$D_0 = D + 4nd\cos\beta'\cos\theta'\cos\alpha' - 4nd\,\Delta\beta'\sin\beta'\cos\theta'\cos\alpha' + p$$

When these four rays are brought to a focus by the telescope T, the complex amplitude A_P at the corresponding point in the interference pattern will be equal to the sum of the complex amplitudes of these four beams. Accordingly, one can write

$$A_P = A\exp(-2\pi i/\lambda)D_0[\exp i(2\delta + \phi - 4\Delta\psi) + \exp\{-i(2\psi - 2\Delta\psi)\} + \exp i(2\psi - 2\Delta\psi) + \exp\{-i(2\delta + \phi - 4\Delta\psi)\}] \quad (21)$$

where A is, as before, the amplitude of a single beam,

$$\phi = (2\pi/\lambda)p,\ \delta = (2\pi/\lambda)\,2nd\,\Delta\beta'\sin\beta'\cos\theta'\cos\alpha',$$
$$\psi = (2\pi/\lambda)\,2nd\cos\beta'\sin\theta'\sin\alpha'$$
$$\Delta\psi = (2\pi/\lambda)\,2nd\,\Delta\beta'\sin\beta'\sin\theta'\sin\alpha'.$$

The intensity I_P at this point is obtained by multiplying A_P by its complex conjugate and is given by the relation

$$I_P = A_P \cdot A_P^*$$
$$= 4I_0\{\cos(2\delta + \phi - 4\Delta\psi) + \cos(2\psi - 2\Delta\psi)\}^2 \quad (22)$$

where $I_0(= A^2)$ is the intensity due to a single beam or, since $\Delta\psi$ is usually negligibly small in comparison with δ,

$$I_P = 4I_0\{\cos(2\delta + \phi) + \cos(2\psi - 2\Delta\psi)\}^2. \quad (23)$$

As can be seen from a comparison of Eqn (23) with Eqn (10), the intensity distribution in the fringe system obtained in this interferometer is similar to that obtained with the double-passed Jamin interferometer described in the previous section, and the same photometric setting can be used with it as well. Since the phase difference introduced between the beams by a given rotation of the detecting element in this instrument is twice as great as that in the instrument described in the preceding section of this paper, the sensitivity of this interferometer should be correspondingly higher. However, this is partly off-set by the fact that β, the angle of incidence of the rays in the horizontal plane on the beam-dividing plate M_1, must be kept lower than the value for maximum sensitivity, if the plate M_1 is to be of reasonable size.

A minor disadvantage of this interferometer is the fact that the maximum permissible angular displacement of M_2 is rather limited. As can be seen from Eqns (17) and (20), when the angular displacement $\Delta\beta$ of M_2 in the horizontal plane is not equal to zero, the expression for the path difference between the AA and BB rays contains a term which depends on θ, the angle of incidence in the vertical plane. Consequently, the extent of the field around the centre ($\theta = 0$) which can be utilized for the photometric setting decreases as $\Delta\beta$ increases. However, this limitation is not serious, since angular displacements of the order of $\pm 10'$ can be measured without any difficulty being experienced on this account.

Acknowledgments

The authors wish to express their thanks to the Director, National Physical Laboratory of India, for permission to publish this paper.

References

HABELL, K. J., and COX, A., 1948, *Engineering Optics* (London: Sir Isaac Pitman and Sons), 234.

HARIHARAN, P., and SEN, D., 1959a, *J. Sci. Instrum.*, **36**, 70.

——, 1959b, *Brit. J. Appl. Phys.*, **10**, 445.

——, 1960, *J. Opt. Soc. Amer.*, **50**, 357.

WOLTER, H., 1955, *Z. Phys.*, **140**, 565.

ZERNIKE, F., 1950, *J. Opt. Soc. Amer.*, **40**, 326.

Interpretation of ultrasonic echo amplitude

by J. C. V. RUMSEY, B.Sc., Grad.Inst.P., United Kingdom Atomic Energy Authority, Development and Engineering Group, Springfields Works, Salwick, Preston

Paper first received 24*th May* 1960, *and in final form* 3*rd October* 1960

Abstract

The equation

$$S = S_0 \int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx$$

is deduced for the amplitude S of the signal received by a transducer, radius a, from a circular discontinuity, radius T, which is plane and parallel to the transducer face at a distance R, where $k = 2\pi/\lambda$ *and* $x = kat/R$, S_0 *being an arbitrary reference amplitude. Experiments verifying the equation are described, and the results quoted. The equation is shown to be more generally applicable than one recently put forward by Krautkrämer. It is shown that the equation obviates the use of the many test-blocks at present used in ultrasonic inspection.*

1. Introduction

ULTRASONIC flaw detection is widely used as a tool for the non-destructive examination of metals, and many descriptions of the construction and operation of equipment suitable for this purpose are published (Heuter and Bolt 1955).

The value of the method is considerably enhanced if a reliable assessment of defect size can be made. In a recent paper Krautkrämer (1959) put forward a method for the determination of the size of defects idealized in the form of flat-bottomed holes, using an ultrasonic reflection technique. In the present paper, a more widely applicable equation will be developed, which leads to a simple evaluation of size in the case of idealized defects. It will be shown that the equation reduces to that of Krautkrämer in the region for which his equation is valid. Confirmation of the equation is obtained from the results of experiments made in the Springfields laboratory and from results quoted by Krautkrämer.

2. Theory

The pressure p, at a point P with Cartesian coordinates (R, t), due to a piston source of radius a situated at the origin of coordinates O and radiating along the R axis, is

$$p = (2\pi\rho u_1 na^2/R)\{J_1(x)/x\} \qquad (1)$$

provided P is in the 'far field' ($R \geqslant a^2/\lambda$) and attenuation is neglected (Vigoureux 1952) (see Fig. 1).

In this expression $x = ka \sin\theta$ ($= kat/R$ for small values of θ), $\theta = \tan^{-1}(t/R)$, $k = 2\pi/\lambda$; $J_1(x)$ is a first-order Bessel function, ρ equals the density of the medium, n is the frequency of vibration and u_1 is the maximum velocity of the transducer face. The ultrasonic intensity at P is given by $I(R, t) = p^2/\rho c$, where c is the velocity of sound in the medium. It follows that the total energy E incident on a

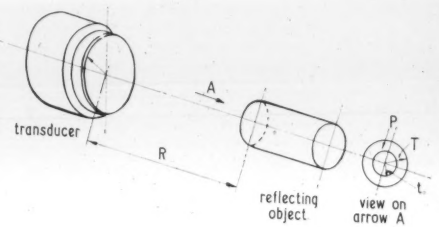


Fig. 1. Arrangement of transducer and reflector.

discontinuity of circular cross section, e.g. a circular disk of radius T perpendicular to the beam with its centre on the beam axis, is

$$E = \frac{4\pi^2\rho u_1^2 n^2 a^4}{cR^2} \int_0^T \left\{\frac{J_1(x)}{x}\right\}^2 2\pi t dt. \qquad (2)$$

Changing the variable of integration from t to x and integrating, the average intensity over the disk I_d is therefore

$$I_d = \frac{8\pi^2\rho u_1^2 n^2 a^2}{k^2 T^2 c} \int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx. \qquad (3)$$

It is assumed that the disk re-radiates as a piston source having a maximum velocity u_2, where u_2^2 is proportional to the average incident energy, i.e.

$$u_2^2 \propto \frac{8\pi^2\rho u_1^2 n^2 a^2}{k^2 T^2 c} \int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx. \qquad (4)$$

This approximation will be less exact as x increases. The ultimate justification for the assumption lies in the experimental verification of the theory which is described below.

To evaluate the energy received back by the transducer, it is necessary to apply Eqn (3) to the disk as radiator, integrating over the surface of the transducer the energy

received from the disk. Thus the intensity at the transducer equals

$$\frac{8\pi^2\rho u_2^2 n^2 T^2}{k^2 a^2 c}\int_0^{kaT/R}\frac{\{J_1(x)\}^2}{x}dx \tag{5}$$

i.e. the intensity is proportional to

$$\left[\int_0^{kaT/R}\frac{\{J_1(x)\}^2}{x}dx\right]^2 \tag{6}$$

and the amplitude is proportional to

$$\int_0^{kaT/R}\frac{\{J_1(x)\}^2}{x}dx. \tag{7}$$

It is of interest to note that Eqn (7) is symmetrical with respect to the crystal radius a and the disk radius T.

The above analysis is strictly applicable to a continuously vibrating piston source only, but experimental results show that it presents a fair approximation to the behaviour of the crystal transducers used for ultrasonic inspection, which generate a pulse of 5–10 cycles duration.

The amplitude of the pulse received by a transducer is indicated by a vertical line of length S on a cathode-ray tube. Assuming linearity of amplification,

$$S = S_0\int_0^{kaT/R}\frac{\{J_1(x)\}^2}{x}dx \tag{8}$$

where S_0 is an arbitrary reference amplitude. Numerical values of the integral are not readily found by analytical methods. The integral is, however, easily evaluated by a graphical method. A graph of $\{J_1(x)\}^2/x$ plotted against x can be drawn and the required integral is the area enclosed between the curve, the x axis and the line $x = kaT/R$. A plot of the integral against kaT/R is shown in Fig. 2. When S_0 has been

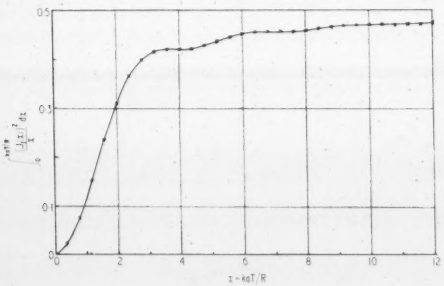


Fig. 2. Plot of $\displaystyle\int_0^{kaT/R}\frac{\{J_1(x)\}^2}{x}dx$ against $x = \frac{kaT}{R}$.

determined, the value of S which will be obtained from a circular discontinuity of given radius at a given distance from the transducer can be calculated, provided the distance is greater than a^2/λ.

The value of S_0 may be found in several ways. One method is to measure the signal amplitude reflected from a circular discontinuity of known diameter at a known distance from the transducer. The appropriate integral may be calculated and S measured, so that S_0 can be obtained by division. An alternative method is to measure the amplitude of the signal reflected from the far side of a parallel-sided section of the medium. All the reflected energy will be received by the transducer, provided attenuation in the medium is low and the section thickness is less than a^2/λ. The signal obtained will correspond to the limiting value (0·5) of the integral as T, and hence x, becomes very large; S_0 is obtained by division as in the previous method. In both methods the reflection coefficient must be the same as that between the medium and the discontinuities, the diameters of which are to be estimated.

3. Experimental verification

Experiments were conducted to verify Eqn (8) over a wide range of values of $x = kaT/R$. The apparatus consisted of a tank 4 ft long, fitted with a carriage which traversed it from end to end. The tank was filled with water and a transducer set up at one end. The plane end of a solid brass cylinder suspended from the carriage was used as a reflector. It was necessary to ensure that the axis of the beam was perpendicular to, and passed through the centre of, the plane end of the cylinder. To accomplish this the following procedure was adopted.

The transducer and cylinder were approximately aligned and the carriage was moved to the end of the tank remote from the transducer (position A). The transducer was rotated about a vertical axis until a maximum reflected signal appeared on the screen. Then the cylinder was rotated about a vertical axis through its plane end until a maximum signal was obtained. The carriage was then moved to a distance from the transducer rather greater than the near field distance a^2/λ (position B), and the cylinder was traversed horizontally across the beam until a maximum signal was observed. The carriage was then returned to position A and the whole procedure was repeated. By this method alignment in a horizontal plane was achieved after a few adjustments. In a similar manner it was possible to obtain alignment in a vertical plane. After satisfactory alignment was obtained, the amplitude of the signal reflected from the end of the cylinder was measured as a function of R, the distance from the transducer. The amplitude was measured by recording the attenuation required in the output signal to reduce the signal height to a fixed level on the screen. Experiments were performed with reflecting cylinders of various diameters and with transducers of various radii and frequencies. It was assumed that the reflection coefficients of the cylinder ends were all identical, so that Eqn (8) applied.

To calculate the amplitude S of the signal reflected from a cylinder of given cross section, it is necessary to take into account the attenuation occurring in the water. Thus

$$S = S_0\left[\int_0^{kaT/R}\frac{\{J_1(x)\}^2}{x}dx\right]\exp(-\alpha R) \tag{9}$$

where α is the amplitude attenuation coefficient in water; S is measured relative to an arbitrary amplitude S_1, so that

$$S/S_1 = (S/S_0)(S_0/S_1) \tag{10}$$

and the attenuation Z inserted in the output signal is given by

$$Z = 20 \log_{10} (S/S_1) = 20 \log_{10} (S/S_0)(S_0/S_1) \quad (11)$$

$$\text{i.e. } \frac{Z}{20} = \log_{10}\left(\frac{S}{S_0}\right) + \log_{10}\left(\frac{S_0}{S_1}\right)$$

$$= \log_{10}\left[\exp(-\alpha R)\int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx\right] + \log_{10}\frac{S_0}{S_1} \quad (12)$$

$$= -\alpha R \log_{10} e + \log_{10}\int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx + \log_{10}\frac{S_0}{S_1} \quad (13)$$

Fig. 3 is a plot of

$$\frac{Z}{20} \text{ against } 0{\cdot}4343\alpha R - \log_{10}\int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx$$

for a typical series of experiments on cylinders of diameter 0·5, 1·0, 1·5 and 2·0 cm at distances R varying from 40 cm (the approximate 'near field' length a^2/λ) to 100 cm, using a 4 Mc/s quartz crystal of radius 1·2 cm; α was taken to be $4{\cdot}1 \times 10^{-3}$ neper/cm. Plotted in this way the points should lie on a straight line of unit slope, according to the above theory. The standard deviation of Z from the best

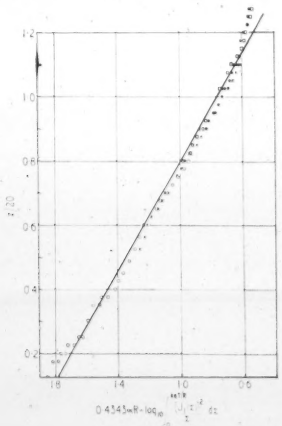


Fig. 3. Plot of $\frac{Z}{20}$ against $0{\cdot}4343\alpha R - \log_{10}\int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx$ for cylinders of various radii using crystal of 4 Mc/s frequency and of 1·2 cm radius. Cylinder radii: ○ = 0·25 cm; × = 0·50 cm; □ = 0·75 cm; ● = 1·00 cm.

straight line through the points is 0·76 dB, and the slope of the line is 0·88 ± 0·02. The deviation of the experimental results from theory may be due to the assumption made in deriving Eqn (4) and to limitations in the accuracy to which attenuation measurements were made (±0·25 dB).

In another series of experiments the signals reflected from the ends of flat-bottomed holes drilled in steel blocks were measured. The holes were of five different diameters in the range 1/32 to 3/32 in. and the ends were at depths from the

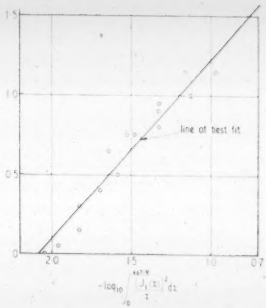


Fig. 4. Plot of $\frac{Z}{20}$ against $-\log_{10}\int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx$ for an experiment using flat-bottomed holes.

surface in the range ½ in. to 3 in. The ultrasonic beam was directed along the axis of each hole in turn, the transducer resting on the surface of the steel. Fig. 4 is a plot of

$$\frac{Z}{20} \text{ against } \left[-\log_{10}\int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx\right]$$

for results obtained using a 5 Mc/s transducer of radius 0·75 cm. In this case, the standard deviation of Z from the best straight line through the points is 2·8 dB, and the slope of the line is 1·13 ± 0·24.

4. Comparison with Krautkrämer's theory

Krautkrämer (1959) gives the equation

$$S/H_0 = k^2a^2T^2/4R^2 \quad (14)$$

for the signal amplitude S from a circular discontinuity relative to a reference amplitude H_0; H_0 is the signal reflected from the far side of a parallel-sided plate, and corresponds to the value of the integral in Eqn (8) when T is very large, i.e. when

$$\int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx = 0{\cdot}5.$$

Thus $$H_0 = 0{\cdot}5S_0 \quad (15)$$

so that $$S/S_0 = S/2H_0 = k^2a^2T^2/8R^2 \quad (16)$$

from Krautkrämer's theory. When $x < 0{\cdot}5$, $J_1(x) \simeq x/2$, so that Eqn (8) reduces to

$$\frac{S}{S_0} \simeq \int_0^{kaT/R} \frac{x}{4} dx = \left[\frac{x^2}{8}\right]_0^{kaT/R} = \frac{k^2a^2T^2}{8R^2}. \quad (17)$$

Thus in the range of kaT/R for which the Krautkrämer equation applies, the signal amplitude predicted by the Eqns (8) and (14) is the same. In Fig. 5 some experimental

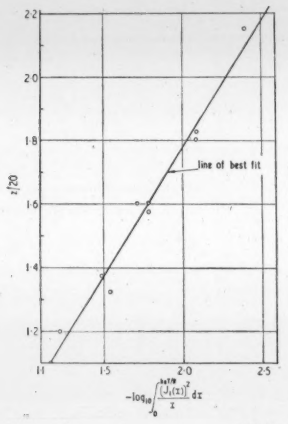


Fig. 5. Plot of $\frac{Z}{20}$ against $-\log_{10}\int_0^{kaT/R}\frac{\{J_1(x)\}^2}{x}dx$ for experimental results quoted by Krautkrämer.

results obtained by Krautkrämer (1959) are shown. The values of x relevant to these results are in the range 0 to 0·7; $Z/20$ is plotted against

$$\left[-\log_{10}\int_0^{kaT/R}\frac{\{J_1(x)\}^2}{x}dx\right],$$

where $Z = 20\log_{10}(S/S_1)$ is the decibel attenuation inserted in the output signal. Z has a standard deviation of 1 dB about the best straight line through the points, the slope of which is $0{\cdot}80 \pm 0{\cdot}09$. The straight line confirms Eqn (8) for small values of x.

In Figs. 6, 7, 8 and 9 are plotted graphs of the percentage

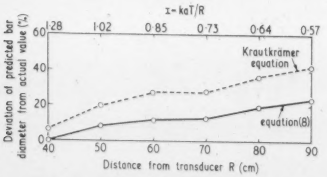


Fig. 6. Comparison of Krautkrämer equation with Eqn (8). Bar diameter, 0·5 cm.

deviation of predicted diameter from actual diameter against distance from transducer. The results are those obtained in the experiment described in § (3) above and plotted in Fig. 3. The bar diameter is calculated using (*a*) Eqn (8), and (*b*) Eqn (14), taking the signal reflected from the 1·0 cm diameter bar at a distance $R = 60$ cm as reference amplitude in each case. The graphs show that for values of x less than about 3·2, Eqn (8) predicts the bar diameter more accurately than does Eqn (14). Eqn (8) is inaccurate for greater values of x because (Fig. 2) there is only a small change in reflected signal for all values of x greater than 3·2. Thus the apparent breakdown of Eqn (8) is due to

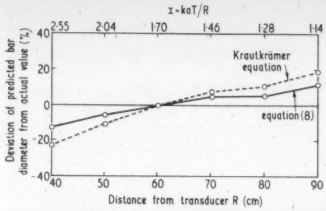


Fig. 7. Comparison of Krautkrämer equation with Eqn (8). Bar diameter, 1·0 cm.

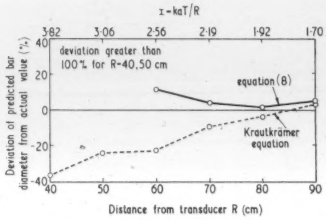


Fig. 8. Comparison of Krautkrämer equation with Eqn (8). Bar diameter, 1·5 cm.

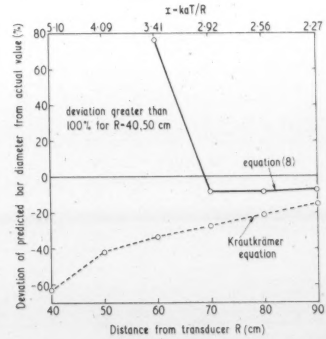


Fig. 9. Comparison of Krautkrämer equation with Eqn (8). Bar diameter, 2·0 cm.

experimental error, rather than to a breakdown in the theory. Values of x greater than 3·2 are of little importance in flaw detection, as it is not usually necessary to determine accurately the dimensions of large flaws.

5. Practical application to defect size assessment

It is current practice to assess the size of defects which have been detected by ultrasonic inspection by comparing the amplitude S of the signal reflected from the defect with

the signal from a flat-bottomed hole in a block of the same material, the end of the hole and the defect being at the same distance from the transducer. This method of assessment involves the construction of a large number of standard flat-bottomed holes, a procedure which is inconvenient and costly.

The equation

$$S = S_0 \int_0^{kaT/R} \frac{\{J_1(x)\}^2}{x} dx$$

eliminates the need for such standards as it predicts the signal amplitude which would be obtained from such holes, provided S_0 is specified initially. It must be emphasized, however, that the equation does not predict the size of defect which gives rise to a given signal, but can only be used to determine the diameter of the flat-bottomed hole which would give rise to a signal of the same amplitude.

The gain control on the testing equipment can be adjusted to give a value of S_0 which has been decided upon by experiment to be optimum for the required inspection. The standard used for reproducing S_0 may be either a flat-bottomed hole or a parallel-sided plate. The advantages of using a parallel-sided plate have been discussed by Krautkrämer. The total energy emitted from a transducer is, in general, much greater than that reflected from a defect, so that for accurate measurement the signal from the far side of the plate must be brought to an arbitrary height on the screen by means of an attenuator inserted in the output, and signals from defects brought to the same level. The use of an attenuator also ensures that the amplifier always operates at the same level, so that errors due to non-linear amplification are avoided.

For routine inspection of materials a graph based on calculations from Eqn (8) can be superimposed on the screen, constructed so that a signal of amplitude greater than the ordinate indicates a defect of rejectable size. The uniformity of the inspection will depend on how uniform are the reflection coefficients of the defects, their shapes and their orientations relative to the transducer face. One method of construction of the graph is to obtain from Eqn (8) the amplitudes $S_1, S_2, \ldots$ of the signals which would be reflected from the disk of size comparable with the smallest rejectable defect at various convenient distances $d_1, d_2 \ldots$ from the transducer. With the transducer set up to receive an echo from the standard (either flat-bottomed hole or parallel-sided plate) used to obtain S_0, attenuation is inserted to obtain the signals, $S_1, S_2, \ldots$ appropriate to the distances $d_1, d_2 \ldots$. The time-base calibration is used to move the signals to positions on the screen at which signals will rise from defects at distances $d_1, d_2 \ldots$ during the inspection of the material under consideration. If tracing paper is placed over the screen the height of the trace in each of the positions can be marked, and hence the required graph is obtained. Using such a graph, the Eqn (8) may safely be applied to routine ultrasonic inspection by semi-skilled personnel.

Acknowledgments

I wish to thank all those who have helped me in this work with encouragement, advice and criticism. I am also grateful to Mr. R. S. Sharpe for lending me the flat-bottomed hole test blocks designed and produced at the Atomic Energy Research Establishment, Harwell.

My acknowledgments are due to Sir William Cook, Director of Development and Engineering, United Kingdom Atomic Energy Authority, for permission to publish this work.

References

HEUTER, T. F., and BOLT, R. H., 1955, *Sonics* (New York: John Wiley and Sons, Inc.).

KRAUTKRÄMER, J., 1959, *Brit. J. Appl. Phys.*, **10**, 240.

VIGOUREUX, P., 1952, *Ultrasonics* (London: Chapman and Hall Ltd.), 5.

CORRESPONDENCE

Measurement of the tensile strength of brittle materials

In their paper published on pp. 281–287 of the June 1959 issue of this *Journal*, Messrs. Berenbaum and Brodie point out that the tensile strength of a brittle material, calculated from the breaking load applied in a bend test, may be more than double the tensile strength found from a tension test. They do not mention the main reason for this apparent discrepancy, which is simply that the conventional formula used to calculate the stress set up by the bending moment is inapplicable if the material does not adhere closely to Hooke's law.

The formula used [authors' symbols, Eqn (5) in the paper] is

$$T = 6G_B/\{(\alpha + \beta)^2 t\}$$

where $\alpha + \beta$ indicates the total depth of the beam d.

This formula is developed in the simple theory of bending on the assumption that the stress–strain law for the material is linear. Cast iron, which the authors include amongst the brittle materials considered, has a stress–strain relation which is approximately linear for small tensile stresses, but as the stress increases the strain increases at a greater rate. In compression the approximately linear relation holds over a greater range of stress. The general shape of the complete stress–strain curve is indicated in Fig. 1.

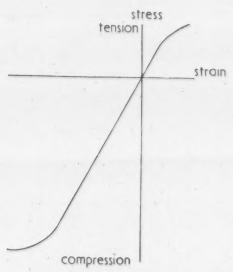


Fig. 1. Stress–strain curve; cast iron.

The distribution of stress over the cross section when the stress–strain relation is linear is shown in Fig. 2. For this distribution, the greatest tensile stress T is given by the formula $T = 6G_B/d^2t$.

The distribution of stress immediately prior to fracture when the stress–strain relation in tension has departed

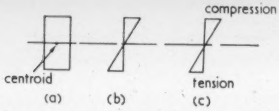


Fig. 2. Distribution of bending stress. Material obeys Hooke's law.
(*a*) Cross section. (*b*) Distribution of strain. (*c*) Distribution of stress.

markedly from the original linearity is as indicated in Fig. 3. The compressive stresses being greater than the tensile, the area in compression is less than that in tension; the position of zero stress or 'neutral axis' has shifted from the centroid towards the compressive face. If the precise shape of the stress–strain curve is known, the position of the neutral axis

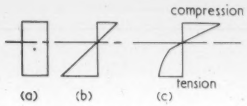


Fig. 3. Distribution of stress prior to fracture.
(*a*) Cross section. (*b*) Distribution of strain. (*c*) Distribution of stress.

can be found by balancing the total tensile and compressive forces. The bending resistance of the section can be found by integrating the moments of these forces. The relationship for the rectangular cross section will be $T = cG_B/d^2t$, where c is a factor and

$$6 > c > 2.$$

This relationship, and hence the amount of the apparent discrepancy under discussion, is dependent on the shape of the cross section. For a circular cross section of diameter d the relationship is

$$T = cG_B/\pi d^3$$

where

$$32 > c > 8.$$

The bend test (three-point loading) or 'transverse' test, as it is called, is widely used for comparing the strengths of specimens of cast iron because of the ease with which specimens can be prepared and tested. The stress calculated from the breaking load by use of the conventional formula is regarded as an indication of strength and is known as the 'modulus of rupture' to distinguish it from the true tensile strength. It is a convenient comparative figure.

The surface finish of a specimen of cast iron is considered to have a minor influence on the results of the bend test. It must be remembered, of course, that owing to the difference in conditions near the surface and at the centre which exists during the solidification of a casting, the properties of the material will vary throughout the bulk. Indeed, these properties will be influenced also by the size of the casting.

The stress conditions obtaining in the transverse test, well established for cast iron, will be valid for other materials with similar stress–strain curves. When the limitations of the test and the difference between 'modulus of rupture' and tensile strength are appreciated, the practical value of the test can be fairly assessed.

Royal Military College of Science, Shrivenham, Swindon, Wilts.

A. ORMEROD

17th November 1959

The authors agree that non-linearity of the stress–strain relation would give rise to the effects described in the paper and probably plays a major part in the case of cast iron. This, however, does not abrogate the general thesis of the paper, that the bending test is extremely sensitive to surface conditions. It should be noted that in the cases of coal and plaster of Paris the stress–strain curves are closely linear up to failure. Non-linearity of the stress–strain relation would, of course, give rise to errors in the other indirect tensile strength tests considered, but this would show up in the validity test suggested, where different sized indenters would give rise to different apparent tensile strengths.

Atomic Power Constructions, Ltd., Research and Development Laboratories, Heston, Middx.

R. BERENBAUM

Westinghouse, Ltd., Elmira, N.Y., U.S.A.

I. BRODIE

9th September 1960

Notes and comments

New Edition of *Notes for Authors*

With the amalgamation of The Institute of Physics and The Physical Society it is desirable to co-ordinate the production of the journals, hitherto published by the two bodies separately.

As is to be expected, the styles employed in the two sets of publications have differed in certain details, and we hope to remove most of these differences with the new volumes just starting. The main alteration which affects the *Journal of Scientific Instruments* and the *British Journal of Applied Physics* is in the form in which references are given. In future this will conform to the Harvard system.

To assist intending authors a revised *Notes for Authors* has been published and is available free of charge from 47 Belgrave Square, London, S.W.1, or 1 Lowther Gardens, Prince Consort Road, London, S.W.7. It contains hints on the preparation of scripts and diagrams, the layout of mathematics and the correction of proofs. Various appendices give lists of preferred spellings, symbols, abbreviations, form of references to be used and a bibliography of reference books and works on technical writing. Intending authors are invited to apply for a copy.

The Institute of Physics and The Physical Society Awards and Appointments

The Council of The Institute of Physics and The Physical Society which recently amalgamated has made the following Awards:

The Duddell Medal to Dr. J. B. Adams, the Director-General of C.E.R.N. (the European Organization for Nuclear Research), for his leadership of that team.

The Charles Vernon Boys Prize to Professor A. W. Merrison of the University of Liverpool for his distinguished research in experimental physics.

The Charles Chree Medal and Prize to Dr. S. E. Forbush of The Carnegie Institution of Washington for his distinguished work on the cosmic radiation.

The Council has also made the following appointments:

Guthrie Lecturer: Dr. D. Shoenberg of the Royal Society Mond Laboratory, Cambridge, has been appointed Guthrie lecturer for 1961.

Thomas Young Orator: Dr. H. H. Hopkins of the Imperial College of Science and Technology has been appointed to deliver the Thomas Young Oration in 1961.

Thermoelectricity Conference, Durham, 1961

An Institute of Physics and Physical Society Conference is being arranged to take place in the Science Department, Durham Colleges, University of Durham, from 10th July to 12th July 1961. Accommodation will be provided in Grey College; meetings will be held in Applebey Lecture Theatre and in the Applied Physics Department.

It is intended that the topics for discussion include thermoelectric refrigeration and generation, and the following headings are intended to indicate the main fields of interests. Basic properties of thermoelectric materials, solid and liquid; methods of determining electrical conductivity, thermal conductivity, and thermoelectric power; properties of materials, problems and techniques associated with device developments; thermoelectric devices and their applications.

Application forms for attending this Conference will be available later from the Institute of Physics and Physical Society. Proposals for contributions should be sent to Professor D. A. Wright, Department of Applied Physics, Durham Colleges.

Mass Spectrometry Conference, Oxford, 1961

The Mass Spectrometry Panel of the Institute of Petroleum in conjunction with A.S.T.M. Committee E–14 are organizing a conference of Mass Spectrometry which is to be held at Oxford University on 12th, 13th, 14th September, and the morning of the 15th, 1961.

Papers dealing with any aspect of Mass Spectrometry will be welcome but special sessions will be held which will deal with high resolution mass spectrometry, the mass spectrometry of organic compounds and the mass spectrometry of inorganic solids.

Authors who wish to present papers at the Conference are asked to send full abstracts by 28th February 1961, to Dr. R. R. Gordon, chairman, Mass Spectrometry Panel, National Coal Board, Coal Research Establishment, Stoke Orchard, Cheltenham, Gloucestershire. The final choice of papers to be presented will be in the hands of a reviewing committee which will make its decisions on the basis of the abstracts submitted.

Details regarding Conference membership and all other arrangements can be obtained from the Organizing Secretary, W. J. Brown, Instrumentation Division, A.E.I. (Manchester) Ltd., Trafford Park, Manchester, 17.

Water Conservation Bibliography

We have received from Price's (Bromborough) Limited, Bromborough Pool, New Ferry, Near Birkenhead, a Bibliography of references in the published literature on water conservation. They have been studying the use of cetyl alcohol for water conservation for a number of years and have built up a classified bibliography covering not only available information on the control of evaporation losses from water storages, but also references in many other fields which are related to the study of the problem.

There are three parts, a subject index, a list of references and an author index. The subject index is divided into sections under the following headings: Evaporation from free water surfaces; Evaporation from soil and vegetation; Monomolecuiar films at the air/water interface; Use of fatty alcohol monolayers for reduction of evaporation from water storages; Hydrological factors associated with water conservation; Meteorological factors associated with water conservation; Other applications for the use of fatty alcohols to reduce the evaporation of water; Properties of the long-chain fatty alcohols; Oil films, multilayers and built-up layers.

Croxson Memorial Lecture 1961

The Society of Non-Destructive Examination announces that the first Croxson Memorial Lecture, in commemoration of the late Mr. Charles Croxson, Founder Chairman of the Society, will be presented on Friday, 17th February 1961, at 6.15 p.m., in the Caxton Hall, Westminster, by Dr. L. Mullins, Ph.D., F.R.P.S., A.I.M., F.Inst.P., on: *The Evolution of Non-Destructive Testing*. Admission to the Lecture will be

by ticket only. Applications for tickets—not more than two per applicant—and for further details of the Lecture, should be addressed to the Honorary Secretary of the Society, Mr. D. T. Carter, E.S.A.B. Ltd., Gillingham, Kent.

Ten-year science plan for UNESCO

A series of recommendations that would govern UNESCO's natural science programme for the next ten years has just been approved by the Programme Commission of the UNESCO General Conference's Eleventh Session now meeting in Paris.

These recommendations give priority for the next decade to activities in three main fields: Co-ordination of scientific activities on both national and international levels; The exploration of the earth's resources and a study of the scientific problems involved in harnessing them; The applications of science and technology for the industralization of developing countries.

The decision of the Programme Commission to approve these priorities will now go before the General Conference itself for a final vote.

Note: Copies of the *Report on the Survey of the Main Trends of Enquiry in the Field of the Natural Sciences* (11C/PRG/5) are available from The Press Division, UNESCO, Place de Fontenoy, Paris 7e.

Progress in Astronautics and Rocketry

Academic Press Inc., 111 Fifth Avenue, New York 3, N.Y., announce that they will publish *Progress in Astronautics and Rocketry*, a new series sponsored by the American Rocket Society. The editor of the series is Profesor Martin Summerfield of Princeton, New Jersey. Volumes will contain selected technical papers offered at the Society's specialized symposia.

Volume 1, dealing with *Solid Propellant Rocket Research*, is edited by Martin Summerfield and based mainly on a symposium held at Princeton, 28th–29th January 1960. Volume 2, entitled *Liquid Rockets and Propellants*, is edited by Loren E. Bollinger, Martin Goldsmith, and Alexis W. Lemmon, Jr. Volumes 3, 4 and 5, due in 1961, will be devoted respectively to *Energy Conversion for Space Power*, and *Space Power Systems*, both edited by Nathan W. Snyder; and *Electrostatic Propulsion*, edited by David B. Langmuir, Ernest Stuhlinger, and J. M. Sellen, Jr.

Nuclear Fusion

We have received the first issue of a new journal published under the above title by the International Atomic Energy Agency. It will appear quarterly and will contain reports of original work and review articles concerning plasma physics and controlled thermonuclear research.

The Scientific Editor is Dr. J. G. Beckerley and there is a Board of Editors consisting of sixteen experts from different countries. The first issue contains five articles including one on Plasma Oscillations by I. B. Bernstein and S. K. Trehan of Princeton University. The abstracts are printed in English, French, Russian and Spanish.

The journal is published by the International Atomic Energy Agency, Kärntner Ring 11, Vienna, Austria, and the annual subscription is 70*s*.

Editorial and publications sales offices

Readers are requested to note that all correspondence of an editorial nature should now be sent to the following address:

The Editor and Deputy Secretary,
The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7 (Kensington 0048).

Correspondence concerning subscriptions, missing journals, and sales of publications should be sent to:

The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1 (Belgravia 6111).

Journal of Scientific Instruments

Contents of the January issue

SPECIAL ARTICLE

Some aspects of research on thin solid films. By K. M. Greenland.

PAPERS

Instrument for recording the dust nuisance emitted by chimneys. By P. A. E. Crosse, D. H. Lucas and W. L. Snowsill.
A platinum resistance thermometer for use at high temperatures. By C. R. Barber and W. W. Blanke.
Automatic correction of counting losses in a Geiger-Müller counter used with periodically varying radiation. By B. B. Trott.
Improved design of spectropolarimeter. By E. J. Gillham and R. J. King.
Modified design for ion-source canals. By K. R. Chapman and G. L. Wrenn.
Automatic recording dilatometer. By C. L. Bell.
Automatic apparatus for measurement of the rate of absorption or evolution of gas. By J. Scanlan and J. R. Dunn.

LABORATORY AND WORKSHOP NOTES

A pumpless circulator and Geiger-counter holder. By R. Scott.
Simple ultrasonic interferometer for student use. By L. Molyneux.
Low-capacitance electrometer earthing key. By L. A. W. Kemp.
Thermoelectrically cooled probe for the determination of semiconductor type. By R. V. Jeanes and K. E. G. Pitt.

NOTES AND NEWS

New instruments, materials and tools. Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.
EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.
REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.
ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611–2.)
SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12*s*. 6*d*. each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.
CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

UNWIN BROTHERS LIMITED, WOKING AND LONDON

Design of scales for industrial instruments

by A. J. MADDOCK, D.Sc., F.Inst.P., M.I.E.E., British Scientific Instrument Research Association, Chislehurst, Kent

MS. received 10th November 1960

Abstract

After discussion of the results of various experimenters on the errors that may arise when reading instrument scales and possible means of reducing them, attention is given to recommendations being prepared by the British Standards Institution. Particularly important is the relationship between size of scale and reading distance and the correlation with this of sizes and spacing of the graduation marks and numerals. An analysis of possible scale ranges and the subdivision of these to proceed in number steps of 1's, 2's or 5's, shows that the number is limited and not all can conform to the desired optimum conditions of design; compromise is therefore necessary particularly in the degree of accuracy to which the scale may be read. The scale designs considered are suitable for industrial instruments having accuracies to the order of 1 to 2%.

1. Introduction

It is the function of a scale, with its associated index, to present to an observer a visual representation of a parameter whose magnitude is given by the intersection of the index with the scale. The position of the index gives an analogue presentation to the observer which he may use directly in estimating the value of the parameter but more often he studies the scale and reads off, and maybe writes down, a numerical value which is in digital form: the observer has thus become an analogue-to-digital converter. For measurement purposes, which may be taken as the true function of an instrument, it is this obtaining a numerical value from the scale and its index which is the important feature. It is essential therefore that any scale should be designed to permit the reading function to be carried out as easily as possible, with the minimum chance of error and to a degree of discrimination in keeping with that of the instrument movement itself.

It has been noted above that for measurement a quantitative reading is required of a desired order of accuracy which entails the conversion process by the observer. However, instruments are often used in a qualitative way only, for check or monitoring purposes, to see that a system is operating at its normal value. Here the scale requirements are less severe and in many cases a subdivided scale could be omitted and the position of the index, in analogue presentation only, would be all that is required. We are all so familiar with a clock, for example, that the time may be ascertained quite well when there are only marks at the five-minute intervals and no numerals: in fact mere dots at the quarters are still satisfactory. In this paper, however, we are concerned with scales for use in instruments from which quantitative readings are to be obtained and these other aspects will be ignored.

Furthermore we shall confine ourselves to industrial type instruments in which the scale is to be read at some distance from the observer and with reasonable times of observation, of the order of some seconds. By industrial type instruments are meant those having an accuracy in the region of 1 to 2%, these figures being taken to be merely indicative of the general quality. Scales for precision instruments, usually studied at close quarters, do not come within the scope of the discussion as special considerations arise in these.

Design data should be such that they will apply to any configuration of scale from straight, through sector, to circular or part circular. The recommendations can therefore apply to linear measuring scales, electrical instruments, pressure gauges, tank level gauges, tachometers, etc.

In reading a scale there is always the possibility of the occurrence of human error, even though the scale itself is perfect in design, and this may be due to a variety of causes. Psychological and physiological processes are involved and scales should therefore be designed to reduce these observational errors as much as possible. To obtain data of this sort it is necessary to experiment with a number of observers and to take large numbers of readings to obtain average data. Much work has been done over many years by various experimentalists. Some of the results quoted may apply only to the particular conditions of the experiment so that some care is needed in interpreting the data. However, enough has been written to enable some generalized information and recommendations to be put forward as a reasoned basis on which scale design should be based. To this end a committee of the British Standards Institution is preparing recommendations for the design of scales for this class of use. This paper, in anticipation of these recommendations, sets out to discuss them and to give some background information that has led up to them. The proposals are, to some extent, a look to the (immediate) future to ensure that new scales are designed rationally: it is certainly true that many existing scales leave much to be desired and many are definitely bad. Some manufacturers have already realized this and improved designs are now evident.

That there is a need for improved presentation and design data for instrument scales was stressed by Murrell (1951) and Golds (1951): examples of various instruments were shown and attention called to the failure of these to accord with good design criteria based upon experimental evidence.

An extensive review of the literature up to 1949 (Princeton University, AF Tech. Rep. 5765, U.S. Air Force, Air Material Command) discussed the various studies that had been made by various workers and was particularly concerned with those pertaining to visual displays in aircraft instruments: a more recent survey was made by Chambers.* In our present paper we shall not discuss the rather special requirements for instrument reading with dark-adapted eyes, since our more immediate concern is with

* Report MRC55/104: HF(IE)(55)8, Medical Research Council and Department of Scientific and Industrial Research.

industrial type instruments under normal illumination. For accurate work high levels of illumination on a dial face are required though this must be without glare; brightness levels of 10 ft-lamberts and over should be aimed at for the dial face.

2. Scale reading habits

It is of interest to consider some of the habits and errors made in reading scales. Those mentioned in this section are concerned with the observer's idiosyncrasies, with difficulties of interpretation, and with handling of numerical data. Difficulties due to sizes of scale markings in relation to reading distance, etc., are considered separately in the next section.

In reading a scale the value indicated by the position of the index on the scale is normally obtained in three steps: Firstly the major division of the scale in which the index lies is noted and reference is made to the numbered major marks above and below the index position. Secondly the minor division in which the index lies is noted and its value determined (this means noting the number of minor divisions between the major marks and deducing from the numerals of the latter what is the value of each minor). Thirdly, if it is desired to read to a subdivision of a minor division, this is done by estimating the position of the index relative to the bounding minor marks and assigning a value to this subdivision after deciding what one unit of subdivision represents. This is quite a complicated procedure and takes place mentally whenever a reading is taken (some observers may work in a slightly different order) unless the particular scale is so well known by repeated use that one or other of these processes can be eliminated.

Scale divisions and numerals attached to the major marks can be arranged in different ways, some of which are better than others. Some scales are straight and some curved, and in extreme cases of the latter may occupy about 270° or even the full 360° of a circle.

Certain observational errors may result from any of the above and an important paper on scale-reading habits by Kappauf (1951a) discusses a number of these and indicates how their incidence or magnitude can be reduced: he summarizes the findings of many workers. The reading habits leading to most difficulty are:

(*a*) Reading by reference only to the *nearer* scale mark to the index, instead of always to the lower. This can result in errors of +1 scale mark value at any of the above three reading processes and it is particularly prone to happen with the numbered major marks, in reversal errors on circular scales when reading in the bottom quadrants and in errors of scale division misinterpretation.

(*b*) The need to add successive components of a reading, as already outlined above. Associated with this can be errors of one numbered scale mark value or of one minor division; reversal errors also arise from this cause. It is responsible for a general excess of positive over negative errors.

(*c*) Rounding out a reading to agree in the total number of digits with other readings taken from the scale. This leads to high error frequencies in regions of a scale just below a major mark which has one more digit than the lower section of a scale, e.g. just below 10 or 100; this habit produces an excess of positive over negative errors.

(*d*) Reading in a left-to-right direction. This is the normal habit transferred from print reading and leads to reversal errors in dial regions where the scale proceeds from right to left, e.g. the lower quadrants of a part-circular or circular scale.

Other reading characteristics leading to errors of interpolation are:

(*e*) Preference for certain numbers when interpolating a division ostensibly into tenths. Here divisions of 0, 2, 8 and sometimes 5 tenths are preferred as is indicated by the frequency of occurrence in Fig. 1 (Yule 1927).

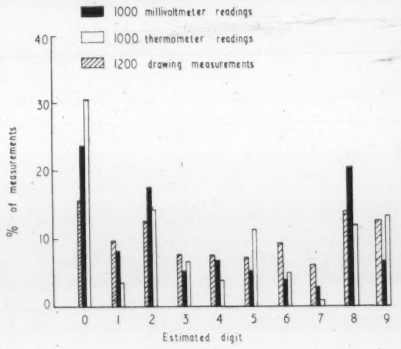


Fig. 1. Number preferences. All numbers had an equal chance of being recorded but observers prefer 0, 2, 8 and sometimes 5 when interpolating. (*After* Yule.)

All 0 to 9 divisions should have appeared equally. This means that it is an easier process to divide into fifths, or halves (perhaps quarters) than into tenths.

(*f*) Number habits, which determine the ease of handling the number system by which readings are determined. Vernon (1946) has shown that divisions numbered in 1's, 2's or 5's give fewer errors than in 4's, 2·5's, 8's (Fig. 2) and the use of 4's for minor divisions or subdivisions of minors leads to the more difficult additions

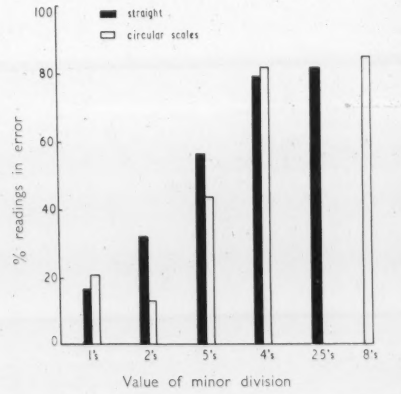


Fig. 2. Observational errors in relation to numbering systems. (*After* Vernon.)

giving values of 4, 8, 12, 16, 20, with resulting greater number of errors.

General recommendations to reduce the above-mentioned errors are given by Kappauf (1951a, b) based upon his own and other experimenters' work. The principal ones are as follows:

(1) Do not use full-circle scales—have a break at the origin and put the zero near the bottom of the dial.
(2) Place numerals and elongated major marks on that side of the scale remote from the index.
(3) Where possible avoid change in the number of digits in numerals at major marks.
(4) The scale should increase in value in a left-to-right or upwards direction. In a part-circle scale locate that portion of the scale where critical quantitative readings are to be made in the upper quadrants.
(5) Design the scale so that the minor divisions are subdivided, by visual estimation, into not more than five parts.
(6) Number the major marks in series of 1's, 2's or 5's; subdivide the scale so that minor marks go in series of 1's, 2's or 5's and so that interpolation of minor divisions goes in 1's, 2's or 5's when appropriately subdivided. Never use systems such as 1·5, 3, 4·5 or 2·5, 5, 7·5 and avoid 4, 8, 12.

3. Size of scale and reading distance

It is fairly obvious that scale size must be related in some way to reading distance and that the degree of subdivision of a scale will have an effect on the readability of that scale. In this latter respect it has been established that too many lines can be provided on a scale leading to great difficulty in reading and resulting errors. This is particularly true if the scale marks are made too long, when a 'palisade' effect results. This was pointed out by Sears (1923) but examples still exist of scales with this defect and engineers' rules are particular examples.

The human eye can distinguish neighbouring objects as separate items down to a certain limit of separation and this resolution leads to a minimum angle which two objects may subtend at the eye to ensure that they are readily resolved. In the case of scales we want the scale divisions to be so spaced that they are clearly legible at the appropriate reading distance.

Many workers have shown that relatively few divisions and wider spacing of scale marks lead to improvement in readability of scales and a reduction of observational errors. Most of them have been concerned with determining the minimum spacing for scale marks on a few sizes of scale to be read at say 28 in. or 10 ft and have not pursued the matter to establish the general relationship between scale spacing and reading distance that would follow. Examples are to be found in the work of Grether and Williams (1947) and of Kappauf and Smith (1950). The diagram of Fig. 3 is based on the findings of the latter and it is seen that the results from different sized dials and different scale numbering systems can be grouped and yield a single curve relating scale spacing and number of correct readings. Laurie, McCarthy and Murrell (1951) realized that it was the angular separation which was the important factor and they extended work of this nature to cover a range of reading distances and dial sizes. Their results also could be combined into a single curve which is shown in Fig. 4. All these workers agree that the number of correct readings increases rapidly as the scale spacing is increased, that there is a fairly sharp transition where the observational accuracy comes near to its highest value, and that thereafter there may be only a very slight

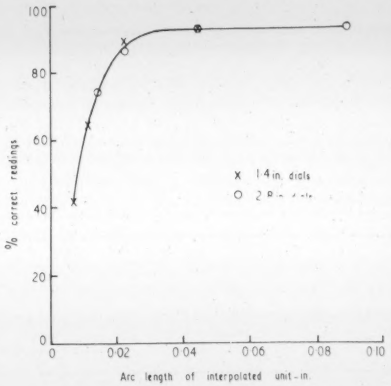


Fig. 3. Observational errors in relation to spacing of graduation marks. Arc length in inches is given for an interpolated one-fifth of a minor division: reading distance 28 in. (*After* Kappauf and Smith.)

improvement with further separation. The high percentage of wrong readings with close spacing may be noted.

Grether and Williams asked their observers to read to $\frac{1}{10}$ of a minor division and used circular scales of 1, $1\frac{7}{8}$, $2\frac{1}{4}$ and 4 in. diameter at 28 in. reading distance. Their critical separation for the estimated interval was found to be 0·05 in. at 28 in. or expressing this in angular measure 0·002 radian. Kappauf and Smith got their observers to read to $\frac{1}{5}$ or $\frac{1}{10}$ of a minor division, depending on the type of scale, and used circular scales of 1·4 and 2·8 in. diameter at 28 in.

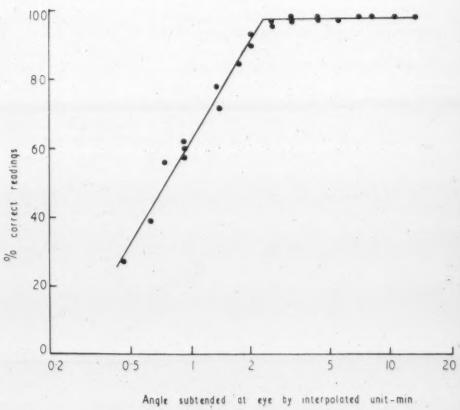


Fig. 4. Observational errors in relation to spacing of graduation marks. Spacing in angular measure for an interpolated one-fifth of a minor division (logarithmic plot of angles). (*After* Laurie, McCarthy and Murrell.)

distance. Critical values are 0·04 in. and 0·08 in. at 28 in. for the two values of interpolation: these are approximately 0·0015 and 0·003 radian respectively. Laurie, McCarthy and Murrell's observers read to $\frac{1}{2}$, $\frac{1}{5}$ or $\frac{1}{4}$ division, depending on the type of scale: part-circle scales ranging from 2 to 8 in. diameter were used over distances from $2\frac{1}{2}$ to 24 ft. They established that by interpolating into fifths of a minor division or less (e.g. quarters) observational accuracies of 95 to 98% were possible with the correct angular separation but that the value was reduced markedly to only 80% for division into tenths. The critical angle, sharply defined, was 2′, i.e. approximately 0·0006 radian. This is only one-third of the American values (interpolation into fifths) and they attribute this to the use of practised observers: over 2000 practice readings were taken on each type of scale before commencing the experiment proper, whereas the subjects used by the American workers seem to have had little or no practice. It is considered (Murrell 1951, 1952) that the British figures give scale sizes more in accord with normal practice and these are the figures which will most likely be included in the B.S.I. recommendations. A practical figure based on the above angular value relating length of an estimated subdivision to 0·08 in. at a reading distance of 12 ft. Expressing this as the relationship between the minimum interval to be read as l in. at D ft reading distance we have

$$D = 150l. \tag{1}$$

The experimental data of the various experimenters have thus given a value for the minimum angle to be subtended at the eye for the smallest division to be read, and further that this latter, estimated by interpolation, should not be less than one-fifth of a graduation interval (minor division) of the scale. From the former we have established the practical relationship between reading distance and minimum division of Eqn (1). We must now determine the total length of the scale in relation to the reading distance.

To fix the overall length of the scale one further factor must be considered, and this is the total number of scale divisions, which must be decided upon in some rational way and not by arbitrary choice. No instrument is without some inherent inaccuracy and it seems logical, therefore, to choose the total number of scale divisions so that the one-fifth of a minor division to which the observer will read does not give so fine a subdivision as to make the reading appear to be much better than that which the instrument is capable of indicating. For industrial purposes it is considered that a scale readable to 1% of full-scale value is generally acceptable and is reasonably related to the general characteristics of this class of instrument. This will correspond to the one-fifth of a minor division so that there will be 100/5 = 20 minor divisions in the total length of scale for optimum design.* Also if L is the scale-base length† in inches, $L = 100l$ and hence the scale-base length and reading distance will be related by

$$D = 1{\cdot}5L. \tag{2}$$

To allow for possible unfavourable conditions of reading, such as lighting which is not ideal, a deterioration in contrast between the scale marks and background of the dial due to ageing or dirt, and to allow for observers whose eyesight is not quite normal, a rather longer scale base is advisable and a suitable relation is

$$D = 1{\cdot}2L \quad \text{or} \quad L = D/1{\cdot}2 \tag{3}$$

L being in inches and D in feet.

A scale designed on this basis can then be read to 1% of its full-scale value (by suitable interpolation) with minimum number of observational errors.

The scale-base length is thus determined, for any reading distance, being based on experimental evidence for the minimum size to make the divisions and a rational choice of total number of divisions in the scale.* To it can then be related the dimensions, proportions and relative spacing of the scale marks, numerals and lettering and other features of the scale and dial: these are given in the later sections.

The scale-base length and all other related features are now the same for any configuration of scale† and Fig. 5 compares the same scale in different forms—straight, sector (90°) and part-circle scale (270°). It is evident from these that quite different dial areas, hence case sizes, are needed to accommodate the same scale: the choice will obviously depend upon the type of instrument and its function.

It is not expected that a manufacturer would make a different size of instrument for each reading distance. Rather the number of case sizes would be limited and the scale-base length in any one will determine the maximum distance at which the scale can be read: this instrument would then be used at somewhat shorter distances until it reached the maximum reading distance of the next smaller instrument and so on. The theoretical continuous relationship between reading distance and scale-base length (Eqn (3)), and from it the dial sizes for a given configuration, would be replaced by a step function in the practical case: this is illustrated diagrammatically in Fig. 6.

4. Scale ranges and subdivision

We have indicated that all graduations and numbering systems on a scale should be such that they proceed in 1's, 2's or 5's (or multiples or submultiples of these by 10) and that the optimum number of interpolated divisions for a scale reading to 1% is 100, resulting in 20 minor divisions if these are interpolated into fifths, or 25 divisions if interpolated into quarters.

* It may be more appropriate on some scales to divide the minor divisions into 4 parts, by eye, to ensure that these estimated divisions are in one of the preferred number series 1, 2 or 5 and examples of this will be found in section 4. In this case there would be 100/4 = 25 minor divisions for optimum design, each estimated division still being 1% of full scale.

† The scale base is the line, actual or implied, running from end to end of the scale, which defines or corresponds to the index path (British Standards Institution 1955).

* By similar reasoning, if a lower class of instrument were being considered and reading to say 5% of full scale (again by interpolation) was of the right order, then only four minor divisions would be needed and the relationship would be $D = 6L$. This would be a very open scale and in fact minor and major divisions would coincide. Kappauf and Smith's work (1950) extended to such 'open' dials and they still fitted into the general pattern of results.

† The evidence that straight, sector and circular scales are all, more or less, equally readable can only be obtained in a general way: no direct results have been reported for quantitative type instruments when readings are taken in the observer's own time. A frequently quoted experiment is that of Sleight (1948) in which he compared these with counter indicators: he found the latter superior and the round dial somewhat better than the straight scales. But the observers were only allowed observation of the scales for a mere flash of 0·12 sec and this has a great bearing on the results; longer times of observation always reduce the reading errors. Grether's (1949) results on several types of altimeter (where the addition of thousands of feet to the scale indication was to be made and this had some bearing on the results) showed little difference between straight vertical and circular scales. It is considered by Murrell (1951) that the design of the scale is likely to have far greater influence on the observational accuracy and time of reading than will its shape.

Table 1 forms the basis for analysis of all possible scale ranges that might accord with these criteria (A. J. Maddock 1959, private communication to British Standards Institution).

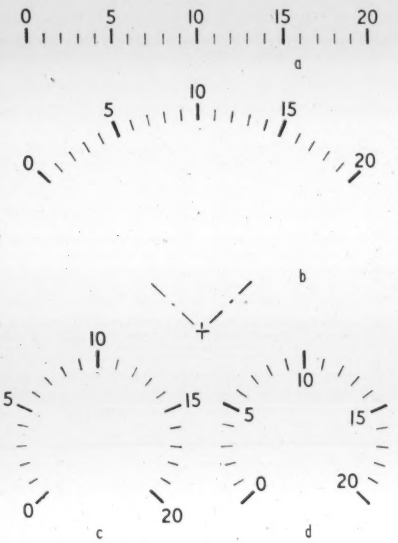


Fig. 5. Areas occupied by the same scale in different forms. *a*, straight; *b*, sector (90°); *c*, part-circle (270°); *d*, part-circle (270°).

It is built up by starting with the total number of estimated divisions on both sides of the optimum of 100 and proceeding by the steps indicated under the Table to list out the full-scale value of all possible ranges. It is apparent from this table that there are only three possible scale ranges that can conform to the optimum, those having full-scale values of 10, 20 and 50, and that all other ranges depart from this optimum. Picking out from this table the only possible scales we have

Optimum	reading to 1%	10, 20, 50
More open	reading to 1·25%	8 (or 80), 40
More open	reading to 1·33%	15
Very open	reading to 1·67%	6 (or 60), 30
More closed	reading to 0·84%	12, 60
More closed	reading to 0·80%	25
Very closed	reading to 0·67%	30

Diagrams of all these scales are given in Fig. 7.

We may note that for the 30, 40 and 50 scales, interpolation of the minor divisions into quarters and not fifths is the appropriate procedure. Taking possible full-scale values between 10 and 100 it is preferable that these should be multiples of 10 and from these, scales in the neighbouring decades follow by multiplying or dividing by 10. We thus see that we can pick out scales of 10, 20, 40, 50, 80 reading to 1 or 1·25% but to find a 30 or 60 scale it is necessary to use the very open scales and read to 1·67%, rather far from the ideal. We do not, however, choose the more closed-in scales to obtain these ranges, especially that for 30 since the divisions would be well below the resolution given by Eqn (2), even using up the full safety factor of Eqn (3): the 30 and 60 scales are therefore difficult to devise and do not fit the scheme well but they are the only possible designs.

An intermediate between 10 and 20 is wanted to avoid too

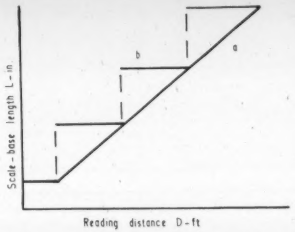


Fig. 6. Illustration of relation between reading distance and scale-base length: (*a*) theoretical continuous relation; (*b*) practical, to limit number of case sizes required.

big a step between ranges here: 15 seems the best choice, reading to 1·33%, although there will be few major marks, as on the 30 and 60 scales. We do not consider 16 since its major marks would proceed in 4's whilst 12 would also proceed in 4's and is not sufficiently intermediate between 10 and 20.

We thus arrive at the following scale ranges as the only possible ones:

10, 15, 20, 30, 40, 50, 60, 80 and start again at 100.

The illustrations of these in Fig. 7 show also the actual value of the estimated division and its percentage value of full scale. These scales should be adequate to cover all needs as there is a good margin of overlap between them (in no case is the ratio between ranges more than 1·5 and more often is less). Several form separate series of 1 : 2 in range, thus 10, 20,

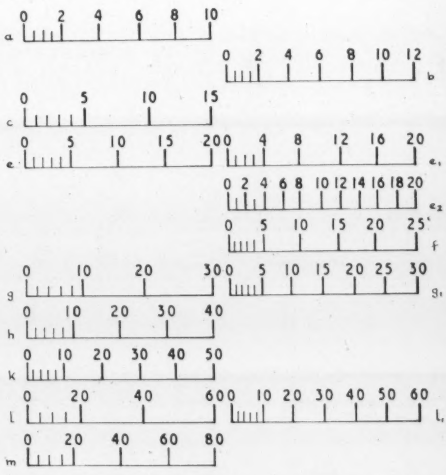


Fig. 7. Diagrams of scales analysed.

Table 1. *Analysis of possible scale ranges*

a. No. of divns in full scale—est. by eye	60			75			80			90		100			110		120			125		150
b. Proportion of full scale %	1·67			1·33			1·25			1·11		1·00			0·91		0·84			0·80		0·67
c. Fraction of minor divn for est. value	⅓	¼		⅓	¼		⅓	¼		⅓		⅓	¼		⅓		⅓	¼		⅓		⅓
d. No. of minor divns	12	15		15	16		20	18		20	25	22			24	30	25			30		
Value of est. divn 0·1																						
Value of minor divn	0·5	0·4		0·5	0·5	0·4	0·5			0·5	0·4	0·5	0·4		0·5		0·5	0·4		0·5		0·5
A. Full scale value	6*	6		7·5	**8**	8	9			**10**	10	11			*12*		12			12·5		15
Value and no. of major divns	2 × 3				2 × 4					2 × 5							2 × 6					
Value of est. divn 0·2																						
Value of minor divn	1·0	0·8		1·0	1·0	0·8	1·0			1·0	0·8	1·0	0·8		1·0		1·0	0·8		1·0		1·0
B. Full-scale value	12	12		15*	16	16	18			**20**	20	22			24		24			*25*		*30*
Value and no. of major divns				5 × 3						5 × 4										5 × 5		5 × 6
Value of est. divn 0·5																						
Value of minor divn	2·5	2		2·5	2·5	2	2·5			2·5	2	2·5	2		2·5		2·5	2		2·5		2·5
C. Full-scale value	20	30*		37·5	40	**40**	45			50	**50**	55			60	*60*	62·5			75		
Value and no. of major divns	10 × 3				10 × 4					10 × 5							10 × 6					

Principal usable scale ranges in bold type.
Usable scales but departing rather far from optimum marked by asterisk.
Scales with too many divisions to accord with the conditions laid down in italic type.

The table is built up as follows in the order given:

a gives the number of estimated divisions in the full-scale length; possible values are given at, above and below 100.
b gives the resulting percentage of full-scale value to which the scale will be read.
d is the number of minor divisions in the full-scale length depending on whether minor divisions are divided by eye into 5 or 4 parts (*c*).

In block rows *A*, *B* and *C* values for the estimated divisions of 0·1, 0·2 and 0·5 respectively are taken and in each of these and row *c* is obtained the value of the minor divisions and thence, with the aid of row *d*, the full-scale value. The value and number of major divisions then follow, remembering that they may only proceed in series of 1's, 2's or 5's.

40, 80 and 15, 30, 60 whilst many are mean values of their neighbours. In fact the complete series comes near to having a ratio of $\sqrt{2} = 1{\cdot}4$ between adjacent ranges and this means that alternate ranges go in a ratio of 1 : 2. Such a $\sqrt{2}$ series would run 10, 14, 20, 28, 40, 56, 80 with a break to start again at 100: the actual run of 10, 15, 20, 30, 40, 60, 80 comes very near to this and we can think of these as the preferred ranges capable of covering all requirements for scales on a decimal notation basis (see Fig. 8). It will be noted that the 50 scale is not required in any of these arrangements and so is left out of the preferred series though it is given in Fig. 7 to illustrate one of the three scales that accord with the optimum criteria.

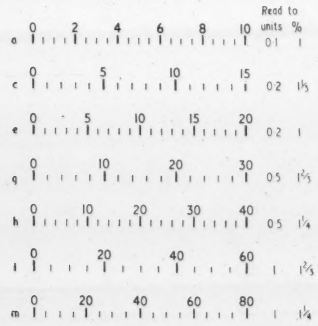


Fig. 8. Preferred scale ranges: same designs when ranges differ by multiples or submultiples of these by 10. Also shown is the minimum unit value to which each scale is read by interpolation and its value as percentage of full scale.

Adoption of these principles and these ranges has the further practical merit of limiting the number of scale arrangements that need be manufactured. Apart from the numerals on the major marks, several scales are identical in dividing and five systems of dividing cover all the possibilities for the series given:

Table 2. *Preferred ranges and divisions*

No. of major divisions	3	3	4	4	5
No. of minor divisions between major marks	4	5	4	5	4
Fig. 8	*l*	*c*, *g*	*m*	*e*, *h*	*a*
Full-scale value	60	15, 30	80	20, 40	10

We should emphasize that the expression of Eqn (3) derived for the optimum 20-division scale should be used in all cases for determining the scale-base length in relation to the reading distance even though, as we have seen above, the smallest subdivision reading may not correspond to 1% of full-scale value. If the length were adjusted to suit the number of subdivisions this would result in different scale-base lengths and hence case sizes, being required for instruments to be read at the same distance, an obviously undesirable feature. It must be accepted that some scales conform to the optimum whilst others have to be more open and will therefore be more easily read; though these latter could be used at a greater distance than the 20-division scale as far as angle subtended at the eye is concerned, yet the dimensions of the scale marks, the numerals and lettering would become too small since their dimensions are all determined to accord with this optimum scale.

Not many scales have possible alternative layouts if one is to keep within the conditions laid down but the 20 (or 200) scale can be arranged in three ways as shown in Fig. 7, each of these scales having the same number of minor graduations and therefore read to the same unit value. These have been discussed by Murrell (1952), whilst Spencer and Dobbie (1954, private communication to British Standards Institution) made a special study of these several arrangements: they also included a fourth system which was the same as that of (e_1) but intermediate values at 2, 6, etc., were emphasized by elongating these to equal the major marks but no numerals were placed thereon. They found sufficiently significant experimental evidence to show that there were less errors made with the 0, 5, 10 . . . system of Fig. 7(*e*) than with any of the others but that the time taken for reading was least for the scale of e_2. The greatest number of errors was made with the scale having the emphasized, unnumbered marks and similar findings are quoted by Kappauf (1951).* Loucks (1944) states that the reading accuracy is increased if less numerals are used. The evidence therefore is in favour of the design of Fig. 7(*e*) and it may be noted that the German standard DIN. 43802 specifies that the scale must not contain more than six numbered marks. In view of this evidence on emphasized intermediate marks these are not incorporated in any of the present scales (e.g. on the 10 or 80).

5. Scale marks, numerals, lettering

5.1 *Scale marks*

Expressing the relative dimensions of the scale marks in relation to the graduation interval, i.e. minor division, the following values have been proposed from time to time:

For the *length of the minor marks* Sears (1923) held that this should not exceed 1 times the interval: Schultz (1925) says that the value should be between $\frac{1}{4}$ and $\frac{1}{2}$ whilst Randle (1946) also favours $\frac{1}{2}$. A value of $\frac{1}{2}$ seems to be a good one to adopt.

For the *length of the major marks* Schultz (1925) proposes a value between $\frac{1}{2}$ and 1 times the interval, i.e. double the length of a minor mark. After tests on a number of ratios 0·85 of the graduation interval, i.e. 1·7 times the minor mark, appears to be the most acceptable.

For the *thickness of the minor marks* Sewig (1936) recommends a value of $\frac{1}{20}$ to $\frac{1}{10}$ of the graduation interval whilst an acceptable *thickness of the major marks* is $1\frac{1}{2}$ times that of the minor (Randle 1946). After tests on a number of values a value of 0·09 relative to the graduation interval for the minor marks and half as much again for the major marks appears to be suitable. If it is desired not to emphasize the major marks by increase of thickness as well as by longer length but to use only the latter, the major marks may be the same thickness as the minors: in this case they will appear to be rather thinner due to their greater length and if it is desired to make them visually of the same thickness about 10% may be added to the width. Another method of emphasis of the major marks which is widely used in Continental practice (Sewig 1936, German standard DIN. 43802) is to thicken only the portion extending beyond the length of a minor mark. This is a neat way and has the merit of keeping the graduation intervals all exactly the same size (see Fig. 9).

It is preferred that a boundary line (arc line on circular scales) should not be used but if used then only one, placed coincidentally with the scale base, should be adopted. Scales have a cleaner appearance and are at least equally readable without boundary lines: this, incidentally, has always been normal Continental practice.

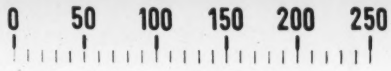


Fig. 9. Example of Continental scale.

Expressing these ratios in terms of our basic dimension, the scale-base length, having the optimum of 20 minor divisions, we arrive at the following actual dimensions per inch of scale-base length, thus making the task of drawing up a scale of any length, to suit any reading distance, an easy one. The dimensions apply to all the scales discussed in Section 4 even though they cannot accord with the optimum in number of minor divisions.

Table 3. *Dimensions of scale marks*

	Per inch of scale-base length	
	Thickness (in.)	*Length (in.)*
Minor mark	0·0045	0·025
Major mark	0·005	0·042
	or 0·0067	
Boundary line (if used)	0·002	

The increased length of major marks should be disposed in a direction away from the index so that the 'inner' ends of all scale marks, major and minor, lie on one line.

5.2 *Numerals*

The subject of numeral design has received extensive attention for instrument dials, for road traffic signs, for car number plates, and for radar plotting rooms. The conditions of reading are obviously different in each of these cases: for example in radar plotting rooms an oblique angle of viewing is involved and for this Mackworth (Bartlett and Mackworth 1950) designed the numerals of Fig. 10(*a*): on car number

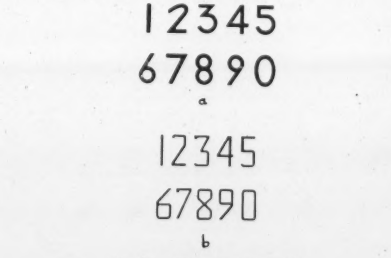


Fig. 10. Design of numerals.
(*a*) Mackworth, (*b*) Berger.

plates all combinations of numerals occur and the time of observation may be extremely short and for this Berger (1944) proposed the angular type of numerals of Fig. 10(*b*). It is considered that, with the type of instrument and conditions of reading with which we are dealing, it is not necessary to adopt the somewhat exagerated designs of these two workers. However, no existing set of numerals fitted all the

* Kappauf did, however, suggest using a V-mark over the middle minor mark if there were ten such marks per major division. This, however, does not apply to scales of the type being considered here where none has more than five such divisions.

characteristics that they should have to accord with the findings of several workers who have studied numeral legibility, so a set has been designed for the British Standards Institution as shown in Fig. 11: this set attempts to have all

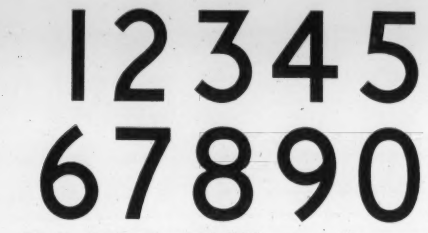


Fig. 11. Design of numerals B.S.I. patterns. (*By courtesy of* British Standards Institution.)

the attributes, to be aesthetically pleasing, to be very legible, to be of such a height-to-width ratio that too much dial space is not occupied, and to be capable of being used for printing or engraving processes. The copyright exists with the B.S.I. but copies are freely available for photographic reproduction; the numerals have also been so designed that they may be constructed with ruler and compasses if it is desired to lay them out on a master.

The general ratio of height to width has been chosen between 1·3 and 1·8 : 1 depending on the actual numeral with an average of about 1·5 : 1 except for the 1, which is obviously different. Recommendations given by other workers are: Berger (1944), 1·9 : 1; Bartlett and Mackworth (1950), 1·4 : 1; Trotter (1916), 1·4 : 1; Soar (1955a), 1·3 : 1.

The stroke width adopted is about 13·5% of height of numeral. Recommendations have been 5% (Trotter 1916); 8–13% (Berger 1944); 14% ($\frac{1}{7}$) (Bartlett and Mackworth 1950); 20% ($\frac{1}{5}$) (Kuntz and Sleight 1950).

Laurie, McCarthy and Murrell (1951) reported that their practised observers were able to read scales with reasonable observational accuracy at distances greater than those at which they could read the numerals on the dial. They performed a separate experiment on the Mackworth design of numerals which they used and confirmed this view. They then established the relationship between numeral height and reading distance at which all-correct readings of the numerals were obtained: a figure taken from this is 0·31 in. at 12 ft distance, i.e. approximately four times the critical length of the interpolated scale division (see section 3).* We have already determined that this is 0·010 in. per inch of scale-base length for the optimum scale (100 estimated divisions) so the numerals work out as 0·040 in. per inch and this is the most acceptable value: it ensures that numeral dimensions and scale mark dimensions are well correlated (this is approximately equal to the length of a major mark). Some variation between 0·035 and 0·045 in. is allowable to cater for cases when there is a large number of numerals on the dial, or a small number, respectively. The spacing of digits forming a numeral of more than one digit is important and these are decided by the designer of the numerals for any combination of digits.

* Craik's (1941) value of 0·312 in. for reading at 23 in. on aircraft instruments (white-on-black) at very low levels of illumination would obviously give numerals far too large if extrapolated to larger reading distances and other conditions of viewing.

Numerals should be so located in relation to the scale that a prolongation of the major mark would pass through the visual or geometric centre of the complete numeral. Numerals should only be placed at major marks and be located on the side of the scale remote from the index whenever possible: an exception would be the part-circle scale where it was desired to reduce the case size as much as possible. In this case the numerals would be 'inside' the scale and some obscuration by the index will occur.

The numerals should always be oriented in the normal viewing position and not tangentially to radial marks (Sewig 1936, Randle 1946). Not more than 3 digits should be used in groups forming the numerals with the exception of the scale maximum value when 4 may be used (Randle 1946).

5.3 *Lettering and other markings*

After the scale, the index and the numerals the next important marking on the dial is the unit designation, that is the indication of the units in which the instrument is reading. The standard abbreviations of BS 1991 should be used (British Standards Institution 1954), e.g. lb/in^2, kw, A, ft, etc., and these should, if possible, be placed outside the area swept by the pointer to avoid any form of distraction: note that many of these are printed in lower-case characters.

In this latter respect also, other marks such as trade marks, maker's name, serial number, test information, etc., should never appear on the main area of the dial: they can conveniently be printed in small type near the edge of the dial so that they are not visible at the normal viewing distance. The general aim throughout should be to keep the area swept by the index clear.

Gill Sans type is suitable for the unit designation and the size chosen should correspond with the numerals in height.

6. Indexes (pointers) and viewing angle

6.1 *Indexes*

In this class of instrument the index has the dual function of indicating a precise location against the scale whilst yet being easily and quickly visible; the eye must be led to the index tip (Vernon 1946, Loucks 1944, Naylor 1954). To comply with the former the actual tip, adjacent to the scale, should have a width about equal to that of a minor mark (Murrell 1951, Randle 1946). The index often overlaps the scale but it should not do so by more than one-third the length of a minor mark. In 'platform' scales, however, there can be no overlap and the index point stops short of the ends of the graduation marks. With this type of scale errors of parallax are eliminated but, unless the gap between the index and scale marks is kept small, another source of error may be introduced due to the necessity for visual projection of the index tip on to the scale.

Circular scales may have taper indexes (pointers) or be parallel-sided with reduced section tip and these fulfil the functions mentioned. The pointer should never project backwards on to the opposing section of the scale: this leads to great confusion in reading and observations on this were noted particularly by Vernon (1946) and Loucks (1944). Since straight scales have not the length of index that a circular scale has, the width of the relatively short index will need to be increased to provide the area necessary for emphasis.

Not much work has been published on indexes and none of it is very conclusive. For example, in circular scales

length-to-width ratios varying between 8 : 1 and 36 : 1 have been quoted (Loucks 1944, Senders, Webb and Baker 1955) as more or less equally acceptable though these represent relatively thick and thin pointers respectively: a value around 10 : 1 might therefore be used as a guide but no more than that. Nowadays the use of spades, crescents and other embellishments on pointers is deprecated as they may lead to confusion and gross errors of reading especially if numerals are hidden by the index.

6.2 *Viewing angle*

The angle at which a scale is viewed and the index separation have obvious bearing on the accuracy of reading a scale and these have been studied by Cohen, Vanderplas and White (1953). Their findings are that with the index in contact with the scale, i.e. no parallax, the errors in reading increase as the cosecant of the angle of viewing decreases from the normal of 90°. There was thus a slow rate of change between 90° and 50° but thereafter the graduation marks appear to come closer together and the errors increase rapidly until at 30° the errors have been doubled. When the pointer is set up from the dial, so that additionally parallax is present, the same general law holds but the angle to which observation could go before rapid change set in was now only 70°, the errors having doubled at 65°. Various spacings between index and scale were used and the above figures apply to normal spacings: the index should therefore be as close to the scale as practicable.

7. Colouring

Maximum contrast is required between the scale marks, numerals and index, and the dial face. For normal usage black markings on a white (or off-white) ground provide this but when viewing by dark adaptation the reverse arrangement would be adopted. Little difference has been found (Kuntz and Sleight 1950, Naylor 1954, Soar 1955b) between these two as far as legibility is concerned if the dimensions of the markings are made thinner in the case of white-on-black. For the normal black-on-white the panel on which an instrument is mounted should have a reflection factor intermediate between the dial and the general surroundings.

The bezel surrounding the dial should not be unduly prominent, should not be of polished metal or highly reflecting paint and in fact might well be of intermediate colour or reflectance between that of the panel and the dial (Murrell 1952). If black is used it should be matt and the bezel should be narrow to avoid any 'heaviness'.

The index will be the same colour as that of the scale markings.

8. Examples

Fig. 8 has illustrated the general layout of the scales: these have been set out on a straight-scale basis. Many modern instruments employ part-circle scales of about 270° angle with the numerals placed inside the graduations, this being done to reduce case size as much as possible. Fig. 12 shows such a scale based on the findings discussed in this paper: it is the 200 scale which is one of those that completely fits all the premises set out. Lastly, in Fig. 13 a comparison between two actual instruments is given to show the difference between the old design of scale with its excess of graduations, double arc lines, extraneous matter on the dial and a modern instrument with scale designed to take account of many of the desirable features which have been discussed above.

9. Summary of principal features

For industrial type instruments or measuring devices, in which the general order of accuracy is to around 1 to 2%,

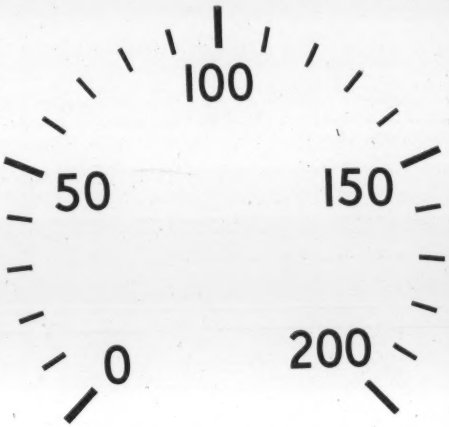


Fig. 12. The optimum design of scale meeting the recommendations laid down (with numerals, however, 'inside' scale to reduce size of case). This scale, as reproduced here, can be read at 8 ft.

the best designs for scales resulting in the greatest legibility and observational accuracy result from the following:

1. *Minor divisions and interpolation*
Do not overcrowd the scale with graduation marks. Observers to subdivide by eye a minor division into not more than five parts: let each of these parts be as near to 1% of full scale as possible, never less. The result is that the optimum total scale length would contain 20 minor divisions. Some small departures from this are necessary to accord with preferred numbering systems (see No. 3 below).

2. *Major divisions*
Arrange the major divisions so that they comprise not more than five minor divisions: in some cases four is correct to establish preferred numbering systems for the minor marks (see No. 3 below).

3. *Numbering systems*
Use numbering systems proceeding only in 1's, 2's or 5's for the major marks, the minor marks, and the estimated subdivisions of the latter: do not use 4's and never 1·5's, 2·5's or 3's.

4. *Position of numerals and major marks*
Place numerals and elongated major marks on that side of the scale remote from the index whenever possible.

5. *Numbered marks*
Number all major marks and do not use emphasized intermediate marks. Do not have more than six major marks in the total length of scale.

6. *Boundary lines*

Do not use boundary lines (arc lines in circular scales). If there is some reason to do so employ one only, never two, and place this on that side of the scale nearest to the index.

7. *Direction of increase in scale value*

Scales to increase in value in a left-to-right, upwards, or clockwise direction, depending on the type of scale. Do not use full-circle scales.

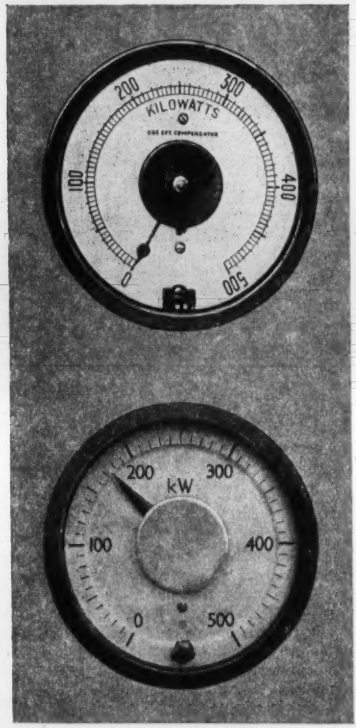


Fig. 13. Examples of actual instrument scales: (*a*) old pattern; (*b*) new pattern. [*By courtesy of* A.E.I. (Manchester) Ltd.]

Though the modern scale shown does not follow all the features discussed, the great improvement is readily apparent.

8. *Reading distance and scale-base length*

Relate total length of scale (L in.) to reading distance (D ft) by the relation

$$D = 1{\cdot}2L \quad \text{or} \quad L = D/1{\cdot}2.$$

9. *Dimensions of marks*

Relate all dimensions of the graduation marks to the scale-base length (derived as in No. 8). This ensures correct proportioning for any reading distance and size of scale. The scale may be arranged in different configurations (e.g. straight, sector, part-circle) but the dimensions are the same for all. Dimensions are:

	Per inch of scale-base length	
	Thickness (in.)	*Length (in.)*
Minor mark	0·0045	0·025
Major mark	0·005 or 0·0067	0·042
Boundary line (if used)	0·002	—

10. *Scale ranges*

Avoid a multiplicity of scale ranges. For decimal notations the following preferred ranges should cover all requirements, each having approximately a $\sqrt{2} = 1{\cdot}4$ ratio to its predecessor: 10, 15, 20, 30, 40, 60, 80 and start again at 100, with multiples and submultiples of these by 10.

These scales are illustrated in Fig. 8 and the resulting economy in design is given in Table 2. They have been built up to incorporate all the above recommendations.

11. *Digits in numerals and units of measurement*

Not more than three digits should be used in groups forming the numerals with the exception of the top mark of the scale when four may be used. For higher range instruments different units should be chosen with the object of reducing the number of digits and to avoid completely the use of multiplying factors which can easily lead to ambiguity: units such as kv, μ in, 1000 ft, assist in this respect.

12. *Design and size of numerals*

Use numerals of B.S.I. design illustrated in Fig. 11, of height 0·040 in. per inch of scale-base length. Variation in this value between 0·035 and 0·045 in. allowable depending on the number of numerals on the dial.

13. *Unit designation and other markings*

For unit designations use the abbreviations of BS 1991, lettering in a sans serif type and of height equal to that of the numerals. Place any such lettering clear of the area swept by the index.

All other matter to be excluded from the dial face but essentials may be printed in small type near the edge of the dial so as not to be visible at the normal viewing distance.

14. *Indexes, viewing angle, colouring*

Take due note of the comments given in sections 6 and 7 regarding indexes (pointers), viewing angle, colouring.

10. Acknowledgments

Thanks are due to the Director of the British Standards Institution for permission to quote material likely to be incorporated in one of their publications and material submitted during the Committee stage, and to the Council of the British Scientific Instrument Research Association for permission to publish this paper.

References

BARTLETT, F., and MACKWORTH, N. H., 1950, Air Ministry Publication, 3139B, H.M. Stationery Office.

BERGER, C., 1944, *J. Appl. Psychol.*, **28**, 208 and 336.

BRITISH STANDARDS INSTITUTION, 1954, BS 1991 (London: British Standards Institution).

—— 1955, BS 2643 (London: British Standards Institution).

COHEN, J., VANDERPLAS, J. M., and WHITE, W. J., 1953, *J. Appl. Psychol.*, **37**, 482.

CRAIK, K. J. W., 1941, Rep. FPRC 342, Air Ministry, Flying Personnel Research Committee.

GERMAN STANDARD DIN. 43802.

GOLDS, L. B. S., 1951, *Proc. Instn Elect. Engrs*, **98,** 671.

GRETHER, W. F., 1949, *J. Appl. Psychol.*, **33,** 363.

GRETHER, W. F., and WILLIAMS, A. C., Jnr., 1947, Rep. TSEAA-694-1E, U.S. Air Force, Air Materiel Command.

KAPPAUF, W. E., 1951a, AF Tech. Rep. 6569, U.S. Air Force.

—— 1951b, AF Tech. Rep. 6366, U.S. Air Force, Aero Medical Laboratory.

KAPPAUF, W. E., and SMITH, W. M., 1950, AF Tech. Rep. 5914, Pt 4, U.S. Air Force, Air Materiel Command.

KUNTZ, J. E., and SLEIGHT, R. B., 1950, *Amer. J. Psychol.*, **63,** 567.

LAURIE, W. D., MCCARTHY, C., and MURRELL, H., 1951, Rep. 47, Admiralty, Naval Motion Study Unit.

LOUCKS, R. B., 1944, School of Aviation Medicine, Project 265, Rep. 2, U.S. Army Air Force.

MURRELL, K. F. H., 1951, Rep. 48, Admiralty, Naval Motion Study Unit.

—— 1952, *Instrum. Pract.*, **6,** 225.

NAYLOR, G. F. K., 1954, *Occup. Psychol.*, **28,** 90.

RANDLE, T. P., 1946, Rep. 27, Admiralty, Naval Motion Study Unit.

SCHULZ, H., 1925, *Indust. Psychotech.*, **2,** 5.

SEARS, J. E., 1923, *Dictionary of Applied Physics*, **3,** 629.

SENDERS, J. W., WEBB, I. B., and BAKER, C. A., 1955, *J. Appl. Psychol.*, **39,** 433.

SEWIG, R., 1936, *Z. InstrumKde*, **56,** 349.

SLEIGHT, R. B., 1948, *J. Appl. Psychol.*, **32,** 170.

SOAR, R. S., 1955a., *J. Appl. Psychol.*, **39,** 43.

—— 1955b, *J. Appl. Psychol.*, **39,** 429.

TROTTER, A. P., 1916, *J. Instn Elect. Engrs*, **54,** 273.

VERNON, M. D., 1946, Rep. A.P.U. 49, Medical Research Council Applied Psychology Research Unit.

YULE, G. U., 1927, *J. R. Statist. Soc.*, **90,** 570.

Etching and polishing studies on magnesium oxide single crystals

by T. K. GHOSH*, B.Ch.E., and F. J. P. CLARKE, B.Sc., Ph.D., A.Inst.P., Atomic Energy Research Establishment, Harwell, Berks.

MS. received 22nd September 1960

Abstract

The action of two etchants on magnesium oxide has been studied. Etch pits, pyramids and hillocks may be produced and both etchants dissolve away the crystal surface. Activation energies for dissolution are estimated and the etching processes discussed with particular reference to dislocations and impurities.

1. Introduction

THE two etchants used in this investigation have been used by other workers (Stokes *et al.* 1958, 1959a) to reveal dislocations and to polish crystals. In early experiments using these etchants various etch figures not previously reported were observed by the authors and, since these might have a bearing on their use to study dislocation movements in magnesium oxide, it was thought desirable to study the processes in more detail. The polishing medium may also be used to obtain very thin magnesium oxide plates, say 100 Å thick, for transmission electron microscopy or spectrophotometric investigations in regions of high optical absorption. Hence a knowledge of dissolution rates is important and details of these have been recorded.

2. Experimental details

Solution A: saturated ammonium chloride and concentrated sulphuric acid in equal parts.
Solution B: 88% orthophosphoric acid.

Other proportions of the constituents of solution A have been used, but it was found that equal parts of each gave the best defined etch pits.

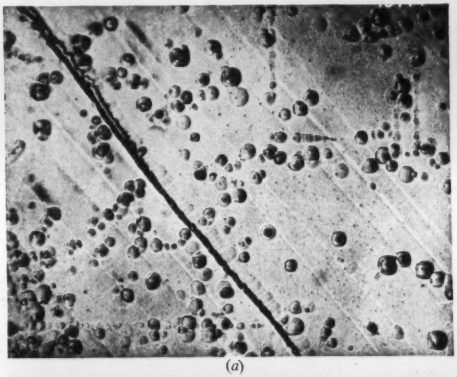
Differing temperatures for etching were obtained by placing the beaker holding the etchant in a Townson Mercer constant temperature bath maintained within $\pm 0{\cdot}1^\circ$ C. To give reproducible results it was necessary to have the etchant continuously stirred.

The thicknesses of material removed by various treatments were measured on a Sigma comparator (Sigma Instrument Co. Ltd., Letchworth, Herts.) in a constant temperature room, maintained at $\pm 0{\cdot}1^\circ$ C during the periods of measurement, and were reproducible to 3×10^{-6} in.

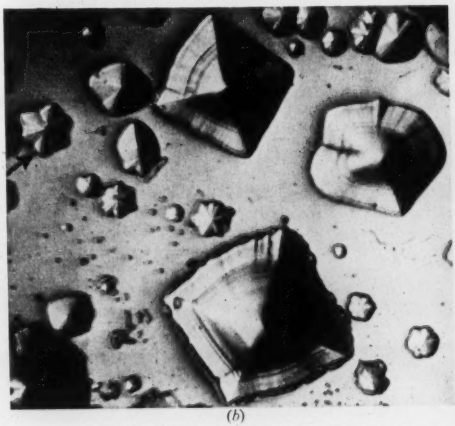
Photographs were taken under various lighting conditions on a Vickers projection microscope. The magnesium oxide used was 99·8% pure (for a typical analysis, see Clarke (1957)).

3. Experimental results

Solution A. It has been established by Stokes, Johnston and Li (1958, 1959a) that this etchant (in slightly different concentrations of component chemicals) reveals dislocations and this has been confirmed during the present work.



(*a*)

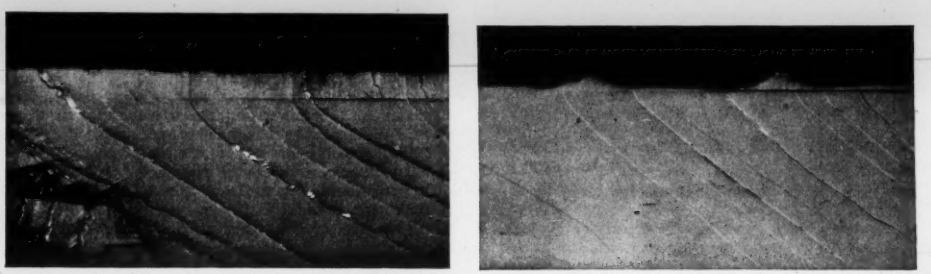


(*b*)

Fig. 1. Typical etch patterns produced by solution A.
(*a*) Etched in solution A for 1 min at 68° C.
(*b*) Etched in solution A for ½ min at 110° C.

* On attachment from Atomic Energy Establishment, Trombay, India.

The size and shape of etch pits are temperature dependent and, at the higher temperatures of etching, hillocks are produced on the surface which are not associated with pits. Typical variations of pits and hillocks are shown in Figs 1(*a*) and (*b*) after immersion in solution A for times of one minute and half a minute at 68 and 110°C respectively. If such etched crystals are now cleaved along a plane orthogonal to that of Fig. 1, wherever this new plane cuts through an etch pit or



(*a*) (*b*)

Fig. 2. Profiles of etch pits and pyramids produced by solution A.
(*a*) Etched in solution A for 1 min at 68° C and cleaved to reveal two pit profiles.
(*b*) Etched in solution A for ½ min at 110° C and cleaved to reveal pyramid profile.

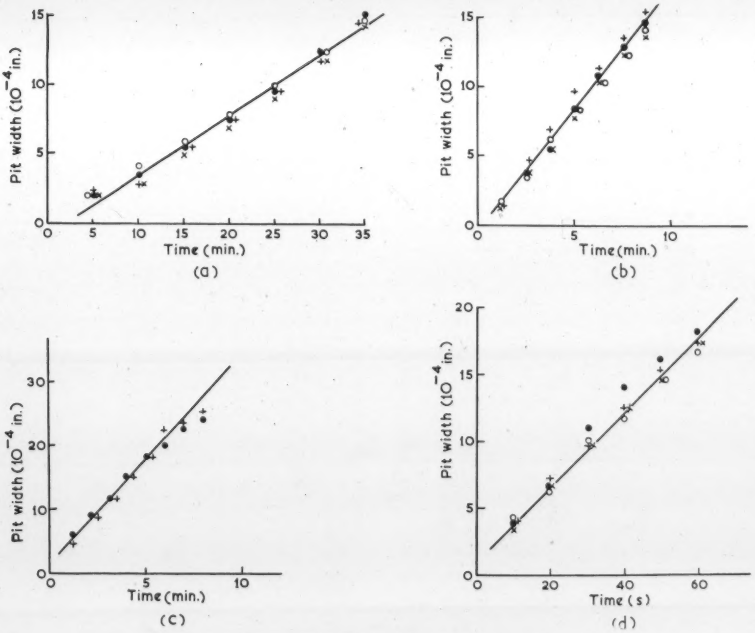


Fig. 3. Increase in pit width with time of etching in solution A at various temperatures.
(*a*) 30° C. Slope, $4{\cdot}35 \times 10^{-5}$ in./min.
(*b*) 50° C. Slope, $1{\cdot}9 \times 10^{-4}$ in./min.
(*c*) 65° C. Slope, $3{\cdot}36 \times 10^{-4}$ in./min.
(*d*) 95° C. Slope, $1{\cdot}66 \times 10^{-3}$ in./min.

hillock the profile is revealed. Typical profiles corresponding to the treatments appropriate to Figs. 1(*a*) and (*b*) are given in Figs. 2(*a*) and (*b*).

The increase in width of the pits with time of etching is linear up to 95° C at least, and Fig. 3(*a–d*) shows graphs of

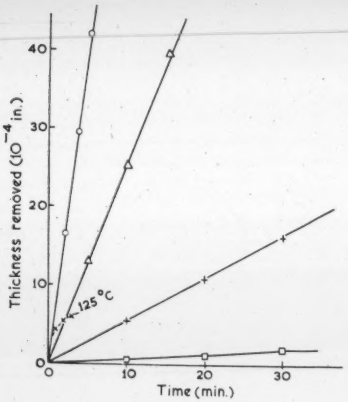


Fig. 4. Removal of magnesium oxide surface in solution A at different temperatures as measured with a comparator.

□ = 25° C, $0{\cdot}3 \times 10^{-5}$ in./min.
+ = 50° C, $2{\cdot}7 \times 10^{-5}$ in./min.
△ = 75° C, $1{\cdot}3 \times 10^{-5}$ in./min.
○ = 95° C, $4{\cdot}0 \times 10^{-4}$ in./min.
× = 125° C, from 5×10^{-4} to $2{\cdot}5 \times 10^{-5}$ in./min.

width of pits against time for temperatures of 30, 50, 70 and 95° C. The ratio of pit width to depth is difficult to determine accurately, but is approximately fifteen for temperatures up to 95° C. This ratio probably increases slightly with time.

This solution sometimes produces pyramids identical in shape to those produced by solution B at all temperatures, as described in the next sub-section. This process of pyramid formation in both solutions is being separately investigated at the present time, and it appears that the production of such pyramids depends upon the crystal used and is associated with impurities. The pits, however, seem to be produced by solution A on all crystals.

Macroscopically the etchant dissolves away the surface, and Fig. 4 is a graph showing thickness dissolved away against time of etching, the thickness measurement being made on a Sigma comparator as already described. Measurements could not be made for etching times beyond a few minutes at the higher temperatures, since the surface of the crystal then became very uneven, with large pit formation. The etching rates at various temperatures are given in Table 1

Table 1. *Etching rates at various temperatures for both solutions*

Temperature of solution (°C)	25	50	75	95	125
Etching rate (10^{-5} in./min)					
Solution *A*	0·3	2·7	13	40	40–2·5
Solution *B*	0·6	2·6	7·7	17·5	17·8

for both etchants A and B. The rate of dissolution increases with temperature, and is constant at any given temperature up to 95° C, but somewhere between 95 and 125° C the rate starts to vary with etching time as shown on the graph for 125° C (Fig. 4).

Solution B. Etch pyramids are produced by this etchant and in contrast to the hillocks produced by etchant A their shape is independent of temperature up to 125° C. Fig. 5 shows photographs of the etch pyramids. The top part gives a plan view of the etch face as revealed by incident light illumination. The lower part shows an orthogonal plane, cleaved after etching to reveal the pyramid profiles. The pyramids at the bottom edge of the top photograph correspond to those of the bottom photograph, but viewed in plan; corresponding pyramids are joined by arrowed lines.

The change in width of the pyramids with time of etching at any temperature is not constant. A typical pattern is that the pyramids initially increase in size to a maximum and then decrease and finally vanish, leaving a small depression in the surface. The actual position of the pyramids on the plane of etching varies very slightly as the surface is gradually

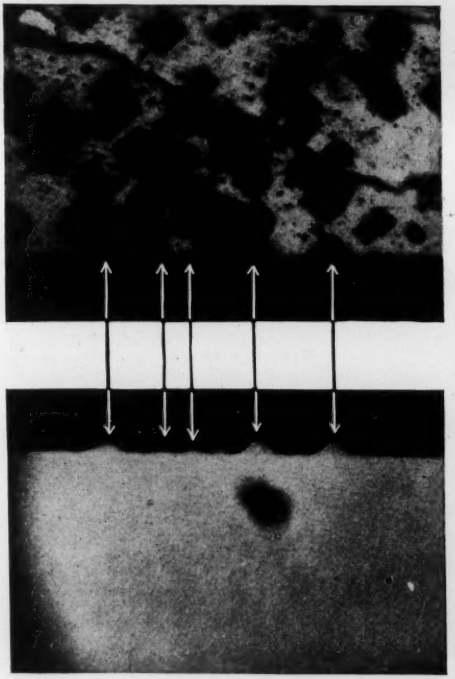


Fig. 5. Pyramids produced by solution B viewed with both plan and side elevation. The crystal has been etched in solution B for 20 min at 60° C, and has then been cleaved. The top photograph gives a plan view of the etched crystal face near the cleaved edge, and the bottom photograph gives a side view of those pyramids cut by the cleavage plane. The plane of this photograph is in the cleavage plane.

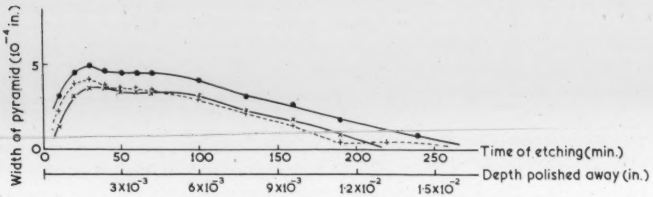
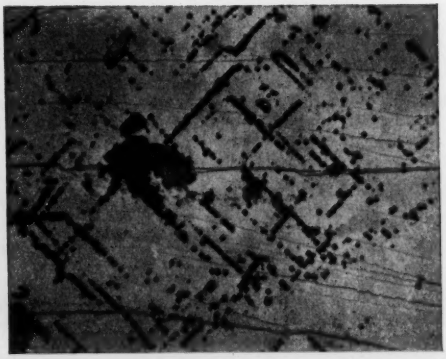


Fig. 6. Treatment in solution B. Width of pyramid is plotted against time of etching and depth polished away.

etched away. The maximum size to which individual pyramids grow has been observed to be between 10^{-4} and 5×10^{-3} in. A typical experiment is represented by Fig. 6, whereon is plotted the pyramid width against time of etching and depth of surface actually polished away. The latter parameter is deduced from the former with the aid of data given later. Apart from occurring randomly over the surface as shown in Fig. 5, pyramids also occur in lines as shown in Fig. 7(*a*). Sometimes, but not invariably, such lines are associated with sub-grain boundaries. Lines of pyramids roughly along ⟨110⟩ directions are also observed as in Fig. 7(*b*), but these do not represent dislocations as revealed by etchant A.

(*a*)



(*b*)

Fig. 7. Differing arrangement of pyramids produced by solution B.
(*a*) Pyramids produced by etching in solution B for 20 min at 60° C.
(*b*) Same treatment as above and pyramids are arrayed along ⟨110⟩.

The individual pits formed by etchant A merge to form a channel along a densely populated slip band (e.g. Clarke 1959). Such a densely packed row of dislocations is sometimes revealed by etchant B after long times of etching. For example, the lines shown in Fig. 8(*a*) only showed up clearly after 40 min at 60° C (i.e. after removal of about 3×10^{-3} in. of surface), and the lines were so shallow that, on cleaving through them as before, no pit profile could be seen at magnifications up to ×2000. Subsequent etching with solution A for 1 min at 70° C showed the typical slip pattern of Fig. 8(*b*). The dislocations revealed on this picture are mainly edge components; however, there are screw components (arrowed) which are not detectable on Fig. 8(*a*). In these experiments no screw components have so far been revealed by etchant B. The absence of pyramids from Fig. 8(*a*) is pure chance and is not associated with the long etching time.

Only one obvious connection has been observed between the pyramids produced by this etchant and the dislocations revealed by both etchants, This is shown in Fig. 9. Such observations are rare and the pattern may result from touching the pyramid tip lightly, thus causing a stress concentration at its base. The authors have not been able to check this unambiguously by deliberate handling because of the fragility of the pyramids. However, the pattern is so like that of rosettes produced by indentation (Stokes *et al.* 1959b) that this explanation is attractive.

Macroscopically solution B also dissolves away the surface and Stokes, Johnson and Li have used it as a chemical polish. Fig. 10 is a graph of thickness dissolved away against time of etching, the thickness measurement being made as before. The rate of dissolution increases with temperature and, in contrast to solution A, is linear at any given temperature up to 125° C, the limit of measurement. The etching rates at various temperatures for both etchants are compared in Table 1. At temperatures above about 130° C the surface dissolves away without pyramid formation.

Identification. Solution A undoubtedly reveals both edge and screw dislocations (Stokes *et al.* 1958, 1959a, Clarke 1959) for temperatures up to 95° C. It is quite clear that when the pits are positioned along slip directions, and can be observed to move and increase in number as the crystal is deformed, then they arise at dislocations. However, it is not so clear

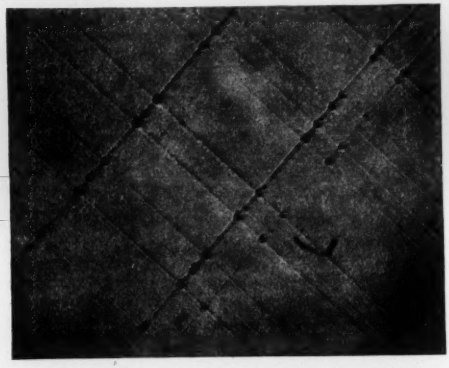
that randomly occurring pits, such as those of Fig. 1(*a*), represent dislocations. Some of the characteristics of these pits are:

(i) as far as has been followed by successive etching (up to distances of 7×10^{-3} in.) their apexes occupy the same relative positions on the crystal surface as the etching proceeds. They do not appear to close, as one would expect of a loop. Gilman and Johnston (1957), for example, found closed loops in lithium fluoride at less than this distance from the surface;

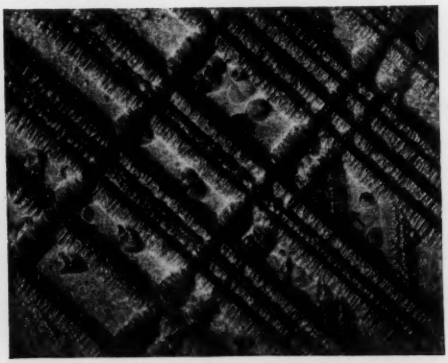
(ii) they are not moved by stresses, nor do stress activated dislocations interact with them;

(iii) solution A will etch both edge and screw dislocations (see Fig. 8(*b*)). If the random pits represented dislocations then, because of (i), they must be pure edge dislocations. Care was taken during the experiments to locate the crystal surface in the same position relative to the microscope, by permanently fixing the crystal to a jig and accurately positioning the jig. This ruled out the interpretation of characteristic (i) that the pits could represent all screw components;

(iv) if pair-faces of a cleaved crystal be compared, only a proportion of the etch pattern appears on both faces.



(*a*)



(*b*)

Fig. 8. Etch patterns revealed by solutions A and B used successively on the same part of the crystal.
(*a*) Solution B, 40 min at 60° C.
(*b*) Solution A, 1 min at 70° C.

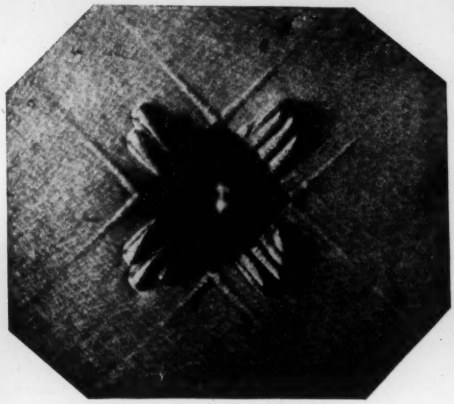


Fig. 9. Photograph illustrating a connection between the pyramids and dislocations. Etch figures revealed by treatment in solution B for 20 min at 60° C.

The authors do not consider it proved, therefore, that *all* etch pits arise at dislocations.

The high-temperature hillocks were not studied in any detail, but it is doubtful if these reveal the position of dislocations, since subsequent etching at lower temperatures shows no correspondence between pits and hillocks. The shell-like shapes arrowed in Fig. 1(*b*) are small hillocks.

Solution B produces very regular pyramidal shapes which are also produced, though not so numerously, by solution A. It has been noticed that these pyramids occur in greater numbers on impure crystals, and on these grounds they are assumed to be associated with impurity deposits in the crystals. The growth of these pyramids reaches a maximum size, and so the process can be followed rather deeper into the crystal than with the random pits of etchant A which quickly overlap, thus making the subsequent following of an individual pit very difficult. Pits and pyramids generally form on different parts of the crystal face, and so they probably represent different crystal imperfections.

4. Discussion of dissolution rates

From Figs. 4 and 10 it is possible, by plotting dissolution rate against reciprocal absolute temperature, to obtain activation energies for the dissolution process and this has been done in Fig. 11. The activation energy for dissolution in respect of solution A is $0{\cdot}65 \pm 0{\cdot}02$ ev and for solution B is $0{\cdot}45 \pm 0{\cdot}03$ ev, both below 95° C.

From the data given, the rates of surface dissolution and etch pit growth can be compared. If one divides the processes between (i) the rate of surface dissolution, R_S before the pits overlap, and (ii) the rate normal to the surface, at

which the pit centre moves into the crystal R_N, then these quantities are related to the apparent rate of increase in pit width R_R by

$$R_N = (R_R + kR_S)/k$$

where k = ratio of pit width to depth, assumed constant $R_N > R_S$ for etch pit growth to occur, and if we take $R_S = 0{\cdot}51 \times 10^{-5}$ in./min, $R_R = 4{\cdot}35 \times 10^{-5}$ in /min, both at 30° C, and $k = 15$ then $R_N = 0{\cdot}8 \times 10^{-5}$ in./min.

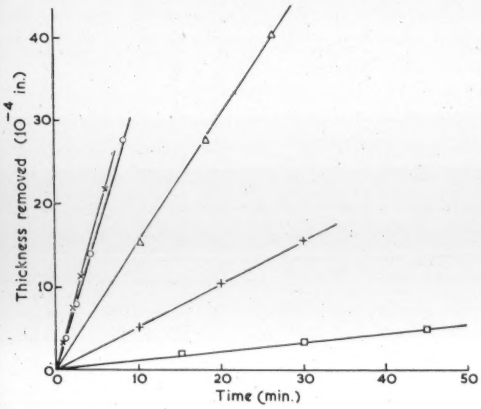


Fig. 10. Removal of magnesium oxide surface in solution B at different temperatures.

□ = 25° C, $0{\cdot}56 \times 10^{-5}$ in./min.
+ = 50° C, $2{\cdot}60 \times 10^{-5}$ in./min.
△ = 75° C, $7{\cdot}70 \times 10^{-5}$ in./min.
○ = 95° C, $1{\cdot}75 \times 10^{-4}$ in./min.
× = 125° C, $1{\cdot}78 \times 10^{-4}$ in./min.

Table 2. *Various etching parameters*

Temperature (° C)	R_S	R_S' (*in./min*)	R_R	R_N
30	$5{\cdot}1 \times 10^{-6}$	—	$4{\cdot}35 \times 10^{-5}$	8×10^{-6}
50	—	$2{\cdot}7 \times 10^{-5}$	$1{\cdot}9 \times 10^{-4}$	—
65	—	$6{\cdot}9 \times 10^{-5}$	$3{\cdot}36 \times 10^{-4}$	—
95	—	$4{\cdot}0 \times 10^{-4}$	$1{\cdot}66 \times 10^{-3}$	—

R_S = rate of surface dissolution, away from etch pits.
R_S' = rate of surface dissolution after etch pits overlap.
R_R = apparent rate of increase in pit width before pit overlapping.
R_N = calculated rate of movement of etch pit apex into crystal and relative to some point fixed in space.

The behaviour of R_R with increasing temperature (Table 2) indicates that R_N will also increase with temperature. Now, if R_S' is the rate of surface dissolution measured after pit overlapping, then since (i) $R_N > R_S$ prior to overlapping, and (ii) R_S' should be more closely controlled by R_N as a surface contour equilibrium is reached, R_S' will be greater than R_S at any temperature. But the linear relationship of Fig. 11 is drawn through points determined largely by the pit-overlapping condition, and so the pre-overlapping points also included might be expected to lie below the straight line, as they do on Fig. 11.

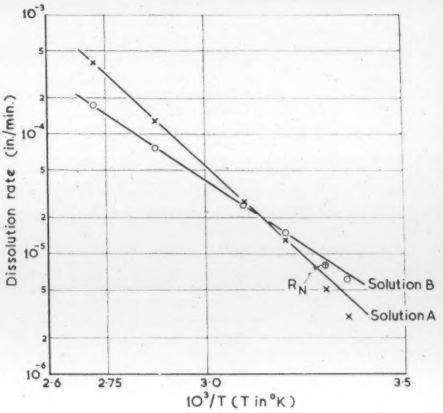


Fig. 11. Plot of log dissolution rate against the reciprocal of temperature for various processes.

If $R_N - R_S' = 0$, one would expect R_N at 30° C to be on the straight line. This does not occur (the exact position of R_N as calculated above is shown on Fig. 11) and would seem to indicate that the lattice disturbances giving rise to $R_N > R_S$ were not present in sufficient density, or were not individually effective over a wide enough surface area, to make R_N control the dissolution process completely.

5. Conclusions

The action of two solutions in etching single crystals of magnesium oxide has been studied.

Solution A: produces pits, hillocks and pyramids. The pits are produced in the temperature range 20–95° C and reveal dislocations. They may also form on random parts of the crystal surface. The hillocks are produced at temperatures above 95° C, but were not studied in great detail. The pyramids are produced in the temperature range 20–125° C and probably form at impurity sites. This solution also polishes away the crystal surface and the activation energy for this process is 0·65 ev.

Solution B: produces pits and pyramids. The pits are indistinct compared with those produced by solution A and reveal edge components of dislocations. The pyramids are identical to those produced by solution A. The solution is primarily a chemical polish for magnesium oxide and the activation energy for this process is 0·45 ev. Dissolution without pyramid formation occurs above 130° C.

Acknowledgments

We are grateful to Mr. J. Williams for reading the manuscript and making several suggestions for its improvement, and to Dr. D. Miles for a discussion on pyramid formation above 130° C.

References

CLARKE, F. J. P., 1957, *Phil. Mag.*, **2**, 607.

CLARKE, F. J. P., 1959, *Special Ceramics Symposium*, 13–15th July 1959 (London: Heywood and Co.).

GILMAN, J. J., and JOHNSTON, W. G., 1957, *Dislocations and Mechanical Properties of Crystals*, 116 (London: Chapman and Hall Ltd.).

STOKES, R. J., JOHNSTON, T. L., and LI, C. H., 1958, *Phil. Mag.*, **3**, 718.

—— 1959a, *Phil. Mag.*, **4**, 137.

—— 1959b, *Trans Amer. Inst. Min. Mech. Engrs*, **215**, 437.

Pulse method for measuring Kerr's constant

by A. PIEKARA and R. KONOPKA, Polish Academy of Science, Institute of Physics, Poznan, Poland

MS. received 26th April 1960

Abstract

A pulse method is proposed for measuring Kerr's effect as produced by a single saw-form electric field pulse of a duration of 5×10^{-5} *sec. The method served for investigating the electro-optical birefringence of liquids whose conductivity was so high as to rule out the use of static methods of measurement.*

Kerr's constant was measured in o-nitrotoluene, m-nitrotoluene and nitrobenzene, these liquids presenting a conductivity ranging from 10^{-7} *to* $4 \times 10^{-11}\ \Omega^{-1}\ \mathrm{cm}^{-1}$.

THE static method for measuring Kerr's constant yields correct results if the medium exhibits low conductivity of the order of $10^{-11}\ \Omega^{-1}\ \mathrm{cm}^{-1}$. If the method is used for liquids of higher conductivity, the results obtained are vastly inaccurate, as the d.c. electric field on the plates of Kerr's cell gives rise to ionic space charges and to a non-linear field gradient between the electrodes. Joule's heat is another factor adversely affecting the accuracy and in some cases rendering measurement impossible.

This is exemplified by the results obtained by Hehlgans (1929) with nitrobenzene of various degrees of purity:

Table 1. *Kerr constant of nitrobenzene by static method (Hehlgans* 1929)

$\lambda = 546\ \mathrm{m}\mu$; $t = 20°$ C

Conductivity ($\Omega^{-1}\ cm^{-1}$)	*Kerr constant* $B \times 10^5$
1×10^{-10}	4·1
2×10^{-8}	2·2

At considerable conductivity, the static method is seen to yield a value of Kerr's constant in nitrobenzene that is by almost 50% lower than the correct one.

To lower the conductivity, the liquids are carefully purified. However, there are dielectric liquids whose conductivity cannot be satisfactorily reduced by any means.

In applying the pulse method for measuring Kerr's constant, the effect of the conductivity is eliminated and correct results are obtained even if the latter exceeds $10^{-7}\ \Omega^{-1}\ \mathrm{cm}^{-1}$.

Hitherto, the pulse method was applied by Benoit (1951) and later by Goodwin (1956), as well as Klages and Küster (1957). These authors used rectangular electric pulses; Benoit obtained them by a mechanical method. The present authors applied saw-tooth pulses, with the voltage rising suddenly and falling linearly during approximately 10^{-5} sec. Thus, Kerr's effect was made accessible to observation on an oscilloscope screen as a function of the variable-field strength. The differences in phase obtained exceeded 180°; at smaller values of the latter, the analyser was initially set at a bias angle, yielding an illuminated field of vision. The method is described in detail as applied to Faraday's effect by Małecki, Surma and Gibalewicz (1957), and by Piekara and above-mentioned co-workers (1957).

The method of measurement is shown in Fig. 1. Voltage

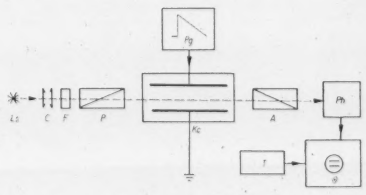


Fig. 1. Block diagram of device.

Ls, luminous source; C, condenser; F, interference filter; P, polarizer; Kc, Kerr cell; A, analyser; Ph, photomultiplier; T, trigger source; O, oscilloscope; Pg, pulse generator.

from a pulse generator is applied to the electrodes of a Kerr cell. The pulse generator works on the principle of discharging an artificial line consisting of two equal LC type elements over a mercury S 15/402 thyratron and the cathode resistance R chosen so that the voltage across it shall fall approximately linearly with the time from a value of U_{max} to $U = 0$. The system was controlled by an EC 50 thyratron.

The planes of polarization of two crossed Glazebrook prisms are set at an angle of 45° with respect to the direction of the field. A d.c. fed 100 w bulb, of the type usually used for projection, provided the luminous source. On passing a condenser and interference filter transmitting a wavelength of $\lambda = 547\ m\mu$, the light beam entered the Kerr cell, to proceed by way of the analyser to a photomultiplier. The variations in light intensity due to electro-optical birefringence were displayed on the double beam oscilloscope screen and registered on a photographic film. The second beam of the oscilloscope simultaneously registered the voltage pulse. Kerr's constant was computed from the formula

$$B = \frac{\beta}{2 \cdot 180^\circ E^2 l}$$

where β denotes the difference in phase between the ordinary and extraordinary beams, in degrees, E the intensity of the pulse electric field in e.s.u. and l the effective length of the Kerr cell in cm.

The method described in the foregoing papers (Małecki *et al.* 1957, Piekara *et al.* 1957) was adapted to measure the phase difference β. The latter was computed from the oscillograms, using the following formulae. If β, the phase difference measured, exceeded 180°, then

$$\beta = n\pi + 2\phi$$
$$\phi = \sin^{-1}(h/h_{max})^{1/2},$$

with h_{max} standing for the maximum deviation of the oscillograph spot corresponding to a phase difference of π or its multiple $n\pi$, and n being an integer (1 in Fig. 2, 2 in Fig. 3, 3 in Fig. 4); h is the deviation corresponding to a phase difference of $2\phi = \beta - n\pi$.

To measure phase differences of less than 180°, a quarter-wave plate, one neutral axis of which was in the direction of the polarizer, was placed before the analyser which was rotated by a known angle Φ previous to measuring the electro-optical birefringence. The difference in phase was computed from the oscillograms (Fig. 5) by the formulae

$$\alpha = \psi + \Phi$$
$$\psi = \sin^{-1}[(h_\psi/h_\Phi)^{1/2} \sin \Phi],$$

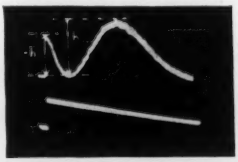


Fig. 3. Oscillogram at phase difference $\beta > 180°$ and $n = 2$.

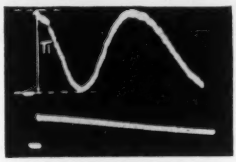


Fig. 4. Oscillogram at phase difference $\beta > 180°$ (= 540°) and $n = 3$.

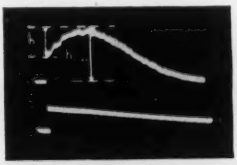


Fig. 2. Oscillogram at phase difference $\beta > 180°$ and $n = 1$.

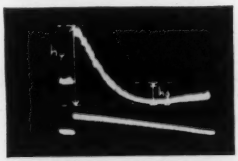


Fig. 5. Oscillogram at phase difference $\beta < 180°$.

Table 2. *Kerr constant of some polar liquids by pulse and static methods*

Liquid	*Conductivity* ($\Omega^{-1}\ cm^{-1}$)	*Kerr constant* $B \times 10^5$	*Author*	*Method*
o-nitrotoluene	6×10^{-8}	2·02	Present authors	pulse
	2×10^{-11}	2·09	Gabler and Sokob (1936)	static
m-nitrotoluene	4×10^{-8}	2·12	Present authors	pulse
	2×10^{-11}	2·16	Gabler and Sokob (1936)	static
nitrobenzene	$4 \times 10^{-11} - 10^{-7}$	3·74	Present authors	pulse
	3×10^{-11}	3·74	Möller (1931)	static
	$< 3 \times 10^{-11}$	3·86	Hehlgans (1932)	static
	$2,5 \times 10^{-12}$	4·03	Gabler and Sokob (1936)	static

where $\alpha = \frac{1}{2}\beta$ is the angle subtended by the direction of the vibration on having passed the quarter-wave plate and that of the polarizer, while h_Φ and h_ψ are the deviations of the oscillograph spot corresponding to the angles Φ and ψ, respectively.

The voltage across Kerr's cell was determined from the oscillograms using the scaling characteristics of the oscilloscope. The effective length l was determined with CS_2, a liquid of known Kerr constant.

The method was used for measuring Kerr's constant in *o*-nitrotoluene, *m*-nitrotoluene and nitrobenzene. The conductivity of the liquids ranged from 10^{-7} to $4 \times 10^{-11}\ \Omega^{-1}\ cm^{-1}$. The results obtained are given in Table 2.

The values of Kerr's constant as obtained for the foregoing liquids are in good agreement with those obtained by Gabler and Sokob (1936) and Hehlgans (1932) by the static method when using liquids of a conductivity of the order of $10^{-11}\ \Omega^{-1}\ cm^{-1}$.

Table 3. *Kerr constant of nitrobenzene of different conductivity by pulse method*

$\lambda = 547\ m\mu$; $t \simeq 20°$ c

Conductivity ($\Omega^{-1}\ cm^{-1}$)	*Kerr constant* $B \times 10^5$
4×10^{-11}	3·72
2×10^{-10}	3·76
6×10^{-9}	3·73
8×10^{-8}	3·70
$1{\cdot}3 \times 10^{-7}$	3·78

To check the effect of conductivity on the results obtained in measuring Kerr's constant by the method described, the latter was measured in nitrobenzene of various degrees of purity. Table 3 gives the results obtained.

The above values of B, as well as the values of B for *o*-nitrotoluene and *m*-nitrotoluene, are the averages from several measurements, the mean deviation amounting to $\pm$ 0·06.

It is seen from Table 3 that, as opposed to the static method (cf. Table 1), the results obtained in measuring Kerr's constant by the pulse method are independent of the conductivity throughout the very wide range of 10^{-11} to $10^{-7}\ \Omega^{-1}\ cm^{-1}$, within the limit of experimental error.

Hence, the present method is adapted for liquids whose conductivity, even following the most careful purification, is still too high to admit of using the static method. The results of such investigations will be published separately.

References

Benoit, H. P., 1951, *Ann. Phys.* (*Paris*), **6**, 561.

Gabler, F., and Sokob, P., 1936, *Z. Tech. Phys.*, **17**, 11.

—— 1936, *Naturwissenschaften*, **24**, 570.

Goodwin, D. W., 1956, *Proc. Phys. Soc.* B, **69**, 61.

Hehlgans, F., 1929, *Z. Phys.*, **30**, 942.

—— 1932, *Z. Phys.*, **33**, 378.

Klages, G., and Küster, R., 1957, *Z. Naturf.*, **12a**, 629.

Małecki, J., Surma, M., and Gibalewicz, J., 1957, *Acta Phys. Polon.*, **16**, 151.

Möller, R., 1931, *Z. Phys.*, **32**, 697.

Piekara, A., Małecki, J., Surma, M., and Gibalewicz, J., 1957, *Proc. Phys. Soc.*, **70**, 432.

Contents

PAGE

Chemistry, and the properties of materials

Applied gamma-ray spectrometry (International series of monographs on analytical chemistry, Vol. 2) . . . 54
Engineering design with rubber 54
Frozen free radicals 55
Graphite and its crystal compounds 55
The proton in chemistry 55
The nature of the chemical bond and the structure of molecules and crystals 56
Some problems of chemical kinetics and reactivity. Vol. 2 . 56
Théorie fondamentale du systeme périodique des éléments . 56
The theory of linear viscoelasticity 56
Unsteady motion of continuous media 56
X-ray powder photography in inorganic chemistry . . 57

Electrical and electronic devices and properties

Advances in electronics and electron physics. Vol. XI . 57
Advances in electronics and electron physics. Vol. XII . 57
The arc discharge. Its application to power control . . 57
Computer engineering 58
Elektronische analogrechner 58
Encyclopaedia on cathode-ray oscilloscopes and their uses . 58
Feedback control system analysis and synthesis . . . 58
Overvoltage research and geophysical applications . . . 59

Heat

Conduction de la chaleur en régime variable 59
Direct conversion of heat to electricity 59
Radiation pyrometry and its underlying principles of radiant heat transfer 60

Mathematical tables

Quantum chemistry integrals and tables 60

PAGE

Nuclear physics, apparatus, techniques and data

Neutron detection 60
Reactor handbook. Vol. 1. Materials 60
The study of elementary particles by the photographic method 61
Pure and applied physics, Vol. 9. Nuclear spectroscopy. Parts A and B 61
Experimental nuclear physics. Vol. 3 61
Atomkraft 62
Nuclear photo-disintegration 62

Optics

The technique of photomicrography 62

Space and aircraft technology

Space technology 62
Leichtbau. Bauelemente, Bemessungen und Konstruktionen von Flugzeugen und anderen Leichtbauwerken . . . 63

Symposia and conference proceedings, etc.

Viscoelasticity; phenomenological aspects 63
Solid state physics in electronics and telecommunications, Vols. 1, 2 and 3 63
Plasma dynamics 63
Structural damping 63
Photochemistry in the liquid and solid states 63
Medical electronics 63
Immediate and low level effects of ionizing radiations . . 64
Electromagnetic wave propagation 64
Gas chromatography abstracts 64

Theoretical physics

The statistical theory of phase transitions 64

New books section

Chemistry and the properties of materials

Applied gamma-ray spectrometry (International series of monographs on analytical chemistry, Vol. 2). Edited by C. E. CROUTHAMEL. (Oxford: Pergamon Press Ltd., 1960.) Pp. xii + 443. Price 50*s*.

Scintillation spectrometers employing thallium activated sodium iodide crystals provide an efficient method of measuring gamma ray energy and intensity, and are consequently extensively used for identifying and assaying gamma emitters.

An accurate interpretation of the scintillation-counter pulse-amplitude distribution requires an understanding of all the effects which determine the shape of the distribution curve and a knowledge of the geometrical factors involved in the measurement.

This book provides the necessary information and enables these factors to be evaluated for any particular spectrometer assembly.

Effects on the pulse amplitude distribution of radioactive decay processes, gamma ray absorption phenomena, and bremsstrahlung radiation are discussed in the first chapter, and effects associated with the counter itself in the second. The fact that pulse amplitude resolution is worse in NaI–Tl than might be expected on statistical grounds has, as yet, received no satisfactory explanation. The current preferences (chiefly those at Argonne) for photomultiplier circuits and other electronic equipment are given, and methods of mounting particularly large NaI–Tl crystals to obtain optimum performance for resolution and background are described with a wealth of practical detail. The use of x-ray proportional counters for low energy gamma ray spectroscopy, to take advantage of the better resolution obtainable compared with scintillators at these energies, is not neglected and is discussed in some detail.

Since each different experimental arrangement requires careful calibration a chapter on this subject is particularly appropriate.

The fourth chapter deals mainly with application to chemical analysis by neutron activation methods using reactors or other neutron sources.

The remainder of the book (pp. 159–439) consists of appendices which include a compilation of pulse amplitude distributions obtained with a 4 in. diameter 4 in. thick NaI–Tl crystal for 133 different nuclides. Tables are included of calculated efficiences for 20 sizes of NaI–Tl crystals and photopeak to total ratios for five crystal sizes. Two other lists give properties of radioactive nuclei, first in order of gamma energy, secondly in order of atomic number.

Whilst a great deal of information similar in character has already been published, it is widely scattered in the open literature, and much is contained in the unclassified reports of the various atomic energy authorities.

The book will be of particular interest to chemists and others who wish to use radioactivation methods of analysis. Scintillation spectroscopy, however, finds many other applications, such as nuclear physics, health physics, ore prospecting, reactor engineering, etc., and workers using the technique in these fields will find much useful information here. One is often faced with the problem of resolving a pulse amplitude distribution into a complex or even continuous gamma ray spectrum. Apart from suggesting simple subtraction of known spectra the book gives little information as to the best procedure in this event, a task which is becoming less tedious as electronic computers are becoming more readily available, but which still remains a formidable problem particularly when using crystals of more modest sizes than 4 in. diameter 4 in. thick.

R. B. OWEN

Engineering design with rubber. By A. R. PAYNE and J. R. SCOTT. (New York: Interscience Publishers; London: MacLaren and Sons, 1960.) Pp. ix + 256. Price 50*s*.

The number of existing books on the important subject of the applications of rubber in engineering is very small indeed, and this new addition is therefore very welcome. Compared with the best-known of existing works, namely *Rubber in Engineering* (H.M.S.O., 1946), the present volume covers a narrower range, but explores it more deeply. The main theme, the dynamic behaviour of rubber, which occupies more than half the total space, is well chosen, not only because of its manifest importance in a wide variety of engineering applications, but also because this branch of the subject is probably the one in which the most extensive developments in systemization and assimilation, to which the authors themselves have made important contributions, have occurred in recent years.

The various methods of representing the dynamic behaviour of rubbers, and in particular the generalized relations between their mechanical response to oscillatory stresses and frequency and temperature are thoroughly developed in Chapter 2, while in Chapter 4 the various types of instruments which have been used for the investigation of these properties are critically examined. Chapter 7 contains an exceptionally clear treatment of the whole question of the transmission of vibrations through rubber members and of the phenomena of resonance; this is particularly relevant to problems of sound insulation and anti-vibration suspensions.

In addition to this main subject there are shorter sections on other physical properties (thermal, electrical, etc.) of rubbers, on test methods other than dynamic, on force-deformation relations in suspension units and other components, on practical design details for components, and on a variety of engineering applications taken from first-hand experience.

The sections dealing with dynamic properties and related test procedures, together with the accompanying scientific background, would be difficult to improve upon. The remainder of the book could be criticized on the ground that there is too much emphasis on detail and not enough on the underlying unifying principles. In this respect the opening chapter, 'The rubber-like state', fails to exploit the possibilities inherent in its title. There is no mention of the derivation of the generalized stress–strain relations of a rubber either from the statistical theory (which involves only a single physical constant, itself determined by the degree of cross-linking) or from the widely used phenomenological theory of Mooney. The elegant solutions to certain basic problems in large-deformation theory (torsion, flexure, etc.) obtained by Rivlin are not discussed in any part of the book. What is lacking, in fact, is any communication of the idea of a theory of *large* elastic deformations, without which the properties of rubbers cannot be properly appreciated. If this had been included the engineer might have been encouraged to take the initiative in attacking problems whose solutions are not to be found in this book.

This omission, and the inclusion of a quite inappropriate introduction, are the only unsatisfactory features of a work which is in other respects outstanding.

L. R. G. TRELOAR

Frozen free radicals. By G. J. MINKOFF. (London and New York: Inter-science Publishers, 1960.) Pp. ix + 148. Price 36*s*.

Comparison of this book with *Formation and Trapping of Free Radicals* edited by Bass and Broida is inevitable. Both were published in 1960 and describe the research completed under the three-year National Bureau of Standards free radical project in the context of previous and contemporary research on unstable species trapped at low temperatures. Bass and Broida's book contains 522 pages, Minkoff's contains 148. Therein lie the advantages and disadvantages of the book under review. It is a thorough survey of the pertinent research which can be read in a few evenings, but it lacks the clarity of the more leisurely treatment in the Bass and Broida book. For example, electron spin resonance is emphasized as a valuable technique in both books. Minkoff's less than twenty pages on electron spin resonance scattered throughout the book can hardly compete with a 42-page chapter by C. K. Jen, who has probably done the most careful and most thorough research to date on the electron spin resonance of frozen free radicals. Minkoff discusses electron spin resonance of molecules in triplet states, which is not treated by Jen, but does not include the first observation of such a resonance by Hutchison and Mangum in 1958. The main advantage of Minkoff's book is the careful organization of the data according to the species studied rather than by the techniques employed.

The layout is particularly fine with detailed outline at the beginning of each chapter, the use of very heavy type for section headings, well reproduced illustrations and thorough author and subject indices. The first 65 pages include chapters on early history, experimental methods, and the basic principles of trapping. An excellent introduction to low temperature techniques is included. The remainder of the book surveys the data obtained about each radical or type of radical. This section occasionally reads like a catalogue or a chapter from the *Annual Review of Physical Chemistry*. Such a presentation is inevitable when so much data is presented in so few pages but makes difficult reading for the non-specialist. The few misprints noticed by the reviewer are obvious and should not lead to confusion.

R. M. DEAL

Graphite and its crystal compounds. By A. R. UBBELOHDE and F. A. LEWIS. (Oxford: Clarendon Press; London: Oxford University Press, 1960.) Pp. vii + 217. Price 35*s*.

The authors have set out to report the current lines of research on graphite. The first chapter is devoted to the crystallographic model for the graphite crystal and the nature of synthetic graphites. It is unfortunate that they have chosen the term 'grain size' for the crystallite diameter since it is frequently used to refer to the size of the original coke particles from which synthetic graphites are made. The model for the near perfect single crystal is adequately treated, but the model for synthetic graphite does not accord with current theory. The existence of long-range order beyond the crystallite diameter has been postulated by a number of workers and has received some experimental confirmation from electron microscopy techniques. These new theories are not presented.

The remaining chapters of the first half of the book set down the properties of the near perfect single crystal and the polycrystalline graphites. The data are collected from the very extensive literature and presented in a coherent form. There is little critical discussion and indeed the authors make it clear in the Preface that to attempt definitive treatment at this stage would be premature. A misprint occurs in Table 9 for the tensile strength of reactor graphites parallel to the grain, which should read 1440 lb/in^2.

The second half of the book records the chemistry of the crystal compounds of graphite. A thorough survey is given of the chemical reactions in which graphite can be involved, and the properties of the resulting compounds are given.

Graphite is finding applications in many fields of activity, and this book can be recommended to the newcomer as a general introduction to the subject. The extensive bibliography of over 1000 references will prove of value to those currently engaged in the subject.

H. H. W. LOSTY

The proton in chemistry. By R. P. BELL. (London: Methuen and Co., 1960.) Pp. vii + 223. Price 42*s*.

The unique position which the proton occupies in chemistry justifies the publication of a book such as this one; for the benefit of physicists it should be pointed out that it deals with protons in solution and not in the form with which they are most familiar. The book will, however, be of interest to anyone concerned with the physical sciences. The author needs no introduction and the subject-matter is presented clearly and in a logical fashion.

The importance of the proton in chemistry is principally due to three factors. Its small size and mass result in its reactions being extremely rapid, while its polarizing power is very high. Secondly, there are signficant isotopic effects in some of these reactions when H is replaced by D or by T. Finally, the existence of acid–base equilibria has important consequences for many aspects of chemistry and allied sciences. These topics are examined thoroughly in the present book. Acid–base equilibria and kinetics are discussed from the viewpoint of the Brönsted theory, with particular emphasis on the effects of structure; in the latter respect, a wide selection of data is carefully assembled and critically analysed. Recent physical methods of measuring directly the rates of acid–base reactions are discussed, by means of which reactions have been studied which occur with practically 100% collision efficiency ($k \sim 10^{11}$ 1 mole^{-1} s^{-1}). A chapter on acid–base catalysis describes recent developments since the publication of the author's earlier book on this subject in 1941.

A number of minor errors appear to have escaped proof-reading. The fifth line from the foot of p. 23 should read Q' (not Q_s); p. 39, line 7 should read $[H_3SO_4^+]$, and the species H_3O^+ is missing from the products in the third equation on p. 40. The labelling of the curves in Fig. 7 should be changed to conform with the text, and the formula for β-naphthol in Table 15 should be $C_{10}H_7OH$. Line 9, p. 135, should read $k_2 \ll k_{-1}$, and the second term on line 8 from the foot of p. 136 should be $[N \equiv \overset{+}{N} . CH_2CO_2^-]$. Line 6, p. 156 should read K_i' and not K_i, the second equation (94), p. 187 should involve XD and not XH, and the third expression on p. 194 relates to Q_v and not Q_v°. The third and sixth columns of Table 26 should be headed E_D and E^*_D/E_D respectively. Some difficulties have also been encountered with the h subscripts on p. 81, where in addition the left hand side of equation 28 should read $-H_0$.

Apart from these minor points, the book is most attractively produced, and the price is reasonable. The reviewer read it with great pleasure and with profit, and it can be heartily recommended.

C. B. AMPHLETT

The nature of the chemical bond and the structure of molecules and crystals. 3rd edn. By L. PAULING. (New York: Cornell University Press; London: Oxford University Press, 1960.) Pp. xx + 644. Price 60*s*.

It is a pleasure to review the third edition of this excellent book. Compared with the earlier editions, it is improved for the student by the addition of a section on the electronic configuration of atoms, the Pauli exclusion principle and its consequences, and the magnetic moments of atoms.

At a more advanced level, several new topics are included in this edition. The author's new theory of the structure of electron-deficient compounds is presented and some applications are discussed. The structure of metals and intermetallic compounds is treated in detail from the point of view of the valence-bond theory. The electro-neutrality principle, that atoms in stable structures tend to have only small amounts of electric charge, is applied to a variety of systems.

The author's theory of resonance is defended against criticisms which have sought to condemn it, in that the structures conceived are idealizations without independent 'real' existence. While admitting that the theory is essentially empirical and qualitative, derived by inductive reasoning, and depending for its successful application on 'chemical feeling', Pauling argues that it is on the same footing as the remainder of chemical structure theory. The successes of the theory in leading to the discovery of new structural features are quoted, and clearly indicate that the status of this theory is at least as great as that of many others of current importance.

The book is very well written, and can be read and understood without specialist knowledge of mathematics.

D. A. WRIGHT

Some problems of chemical kinetics and reactivity. Vol. 2. By N. N. SEMENOV. Translated from the Russian by J. E. S. BRADLEY. (Oxford: Pergamon Press, 1959.) Pp. x + 168. Price 35*s*.

This book was written as an introduction to a symposium on chemical kinetics held in Moscow in 1955. Volume II deals with branched chain reactions and thermal ignition, radical reactions having been dealt with in Volume I.

The author gives a stimulating account of the chemical kinetics of combustion, although the treatment is based largely on the work of the Russian school, and the author, quite correctly, makes no claim to an exhaustive treatment of the subject.

The book is divided under the following chapter headings: Thermal ignition; Chain self-ignition; Chain ignition in hydrogen–oxygen mixtures; Interaction of chains and degenerately branched chain reactions. This approach to combustion, based on the concept of chain reactions, gives coherence to a subject which can otherwise appear difficult and confused; both volumes of this book can be read to advantage by all workers in the field.

Volume II is, unfortunately, not devoid of errors and also contains a large errata to Volume I; it is to be hoped that this deficiency will be corrected in what is otherwise an attractive and most interesting book.

IAN FELLS

Théorie fondamentale du systeme périodique des éléments. By S. DOCKX. (Bruxelles: Office International de Librairie, 1959.) Pp. 182. Price B.f. 300.

It is remarkable how strong a hold the periodicity of the elements has always exerted upon the scientific mind. Associated above all with Mendéléeff's genius, it was at various times—including our own day—the theme of much fundamental thought. All this, the author of the present book brings out well. It is not exactly a story to be read, but to be kept for reference on historical points. The writer is perhaps a little pessimistic about our detailed knowledge of the peculiar structure of the rare earths. The work of Saha and contemporaneous research in the late nineteen-twenties seems to have escaped his net. It fits quite well into his elegant formalism.

An important remark is the one in which he shows magnetic and electronic periodicity to be transposable, by the interchange of the appropriate matrices.

F. I. G. RAWLINS

The theory of linear viscoelasticity. By D. R. BLAND. (Oxford: Pergamon Press, 1960.) Pp. vi + 125. Price 35*s*.

This is a refreshing book, for the author quickly leaves the rather arid and already overworked terrain of one-dimensional spring and dashpot models and proceeds to give a mathematically meticulous but nonetheless readable account of the properties of isotropic viscoelastic solids and liquids based on the (three-dimensional) stress/infinitesimal strain-history relations similar to those first used by Boltzmann for describing the 'elastic after-effect' in solids. Material from the research literature of the past decade is here drawn together for the first time.

Chapter II contains the author's recent derivation of the strain/stress-history relations for an arbitrary 3-dimensional network of springs and dashpots; expressions for the stored and dissipated energies are derived together with certain necessary conditions on the time-derivative operators in the operational form of the stress–strain relations. One wonders, however, in what sense this theory is to be regarded as a valid structural theory: the author states on p. 17 that certain other microscopic structure theories "are mechanically equivalent to a network of springs and dashpots" but it is not clear what is meant by 'mechanically equivalent' in this context nor whether the statement is in fact correct when applied to those of the theories referred to which are based on a model of independent long chain molecules suspended in a viscous environment.

The second half of the book contains a systematic treatment of stress analysis in viscoelastic bodies under the chapter headings 'sinusoidal oscillation problems', 'quasi-static problems', and 'dynamic problems'. Methods of transforming viscoelastic problems into equivalent elastic problems are treated, including Lee's recent method which has led in particular to a solution of the Hertz contact problem for viscoelastic bodies—a problem which is complicated by the occurrence of moving boundaries. In several places the author gives a good succinct account of the relevant properties of typical real materials; the final chapter contains a practical account of the fitting of models to experimental data. The book should prove to be of particular and lasting value to the growing number of physicists, engineers, and rheologists who are concerned with polymeric systems.

A. S. LODGE

Unsteady motion of continuous media. By K. P. STANYUKOVICH. (Oxford: Pergamon Press, 1960.) Pp. xiii + 745. Price £5.

The publishers justify the use of photolithography of the typed manuscript in the printing of this book on the grounds of expense and the speed of reproduction. The format cer-

tainly exposes the reviewer to eye-strain—the symbols in the equations are particularly small and, in places, indistinct—but, aesthetic considerations apart, it is not perhaps so grave a disadvantage if, as presumably the normal reader would, it is taken in small doses.

The book is, the editor writes, completely self-contained and no previous knowledge of gas dynamics is required. . . . "It can be used as a text-book for graduate students and as a work of reference for the many research scientists concerned with the present day problems in gas dynamics." I myself would be unwilling to recommend it in the first category as a text-book to supplant Courant and Friedrich's *Supersonic flow of shock waves* to which the author makes graceful reference. Sheer weight of material, some of it appearing in English for the first time, puts it in the second category though even here I found the presentation heavy going in places (and perhaps even ill-digested in parts). I have also to make a comment, not I am afraid for the first time in similar contexts, that the references are so very sparse as to be almost ludicrous in a book of this size—an author index of 30 names and a bibliography of 49 works. In short I would place this in my second rather than my first line of defensive reference books and note as of special interest to physicists the application to cosmological problems in the last three chapters.

L. HOWARTH

X-ray powder photography in inorganic chemistry. By R. W. M. D'EYE and E. WAIT. (London: Butterworths Scientific Publications, 1960.) Pp. viii + 222. Price 45*s*.

The authors of this useful book are workers at the Atomic Energy Research Establishment, Harwell, and their approach to x-ray crystallographic techniques is through their applications in those studies of inorganic materials that are of importance in the fields of nuclear energy and electronics. Since the available material is often in the form of powder, they have concentrated especially on x-ray powder photography, but in fact much of their book necessarily applies to all x-ray diffraction methods and to crystalline material of any composition and any texture. However, their aim has been to provide, in the form of a conveniently small volume, only that amount of information that a research chemist, with no previous experience of x-ray crystallography, would need to enable him to use this particular technique as a tool. This book does not, therefore, in any way compete with *X-ray diffraction by polycrystalline materials* (The Institute of Physics, London 1955), to which, indeed, it constantly refers the reader who requires more details. My guess would be that research chemists and others needing to identify or to know more about the solid materials they are studying, or wishing to understand non-stoichiometry, phase transitions, the structures of mixed crystals and the mechanism of solid-solid reactions, will find their appetites tremendously whetted by this book, and particularly by Chapter 8 which gives many examples of applications. If they conscientiously work through each chapter, using the references recommended for further study, they will know certainly a great deal of x-ray crystallography by the time they are through.

KATHLEEN LONSDALE

Electrical and electronic devices and properties

Advances in electronics and electron physics. Vol. XI. Edited by L. MARTON. (London and New York: Academic Press, 1959.) Pp. xi + 523. Price £5 7*s*. 6*d*.

This is a well-established series covering a remarkable range of subjects in pure and applied physics. The present volume includes topics from electron production in beta-decay to automatic data processing.

Sternheimer gives a good account of the parity non-conservation in weak interactions, surveying the main experiments in detail but quoting rather than discussing the results of the theory. Photoemission from complex layers is fairly clearly discussed by Gorlich in a language which could have done with one further stage of translation into English: the persevering reader will find it rewarding. Secondary emission is dealt by Hachenberg and Brauer, with a very good account of the theory as well as of experimental knowledge.

Pinsker from Moscow gives a most interesting account of the important Russian work on structure analysis, especially of semiconducting crystals, by electron diffraction. Clark Jones gives a lucid and useful discussion of the quantum efficiency of detectors for visible and infra-red radiation. Unfortunately, the controversy between Clark Jones and Hanbury Brown and Twiss, over the behaviour of grey detectors not in thermal equilibrium with the ambient radiation, was unresolved when the article was written. In consequence, as Clark Jones observes, his expression for the detective quantum efficiency of such detectors is not entirely satisfactory. This is particularly regrettable since the article is concerned mainly with the situation where noise from ambient radiation is not negligible. If, however, the ambient radiation is itself of low intensity, no difficulty arises.

Ribbet's article on radio telemetering is mainly of engineering interest. Konigsberg discusses mainly the uses of operational amplifiers, with a short final section on recent advances in design. The problem here arises in analog computing, where passive networks for analytical operations combine high attenuation with high accuracy, so that good d.c. amplifiers (with as high a working bandwidth as conveniently possible) become very necessary.

Barlow, Ovenstone and Tonemann give a general discussion of automatic data processing in the physical sciences. Such an article, even if it contains little that we ought not to know already, has the function of warning us that we should be thinking of rational methods for handling scientific data while the files are only waist deep, without waiting till the stuff reaches our ears.

G. A. P. WYLLIE

Advances in electronics and electron physics. Vol. XII: Photo-electronic image devices. Edited by J. D. MCGEE and W. L. WILCOCK. (London and New York: Academic Press, 1960.) Pp. xii + 397. Price 86*s*.

This volume is a record of a symposium held in September 1958, at the Imperial College of London University, which is believed to have been the first devoted exclusively to the discussion of photoelectronic image devices. The symposium was convened by Professor J. D. McGee and thirty papers were presented.

The arc discharge. Its application to power control. By H. DE B. KNIGHT. (London: Chapman and Hall Ltd., 1960.) Pp. xix + 444. Price 63*s*.

The title of this book is really rather misleading. It is, in fact, concerned with the theory, design, construction and operation of thyratrons and ignitrons, and not, as might be inferred from the title, with the electric arc except in the rather narrow context of the above-mentioned devices.

In the first of the three main sections of the book the

design and construction of thyratrons and ignitrons is described in some detail, although the treatment is of necessity broad rather than deep. The second section sets out clearly in simple terms enough of the theory of low pressure gas discharges for the newcomer to the subject to understand the operation of the devices with which the book is concerned. The remainder of the book covers the application of thyratrons and ignitrons as circuit elements. This is the longest section, and probably deals with its subject more thoroughly than any other text available.

The book is well produced, the diagrams clear and the number of errors is remarkably small. Users of thyratrons and ignitrons, as well as students and others wishing to increase their knowledge of these devices, will find this a very useful work. M. P. REECE

Computer engineering. Edited by S. A. LEBEDEV. (Oxford: Pergamon Press, 1960.) Pp. vii + 184. Price 40*s*.

Academician Lebedev has edited a small volume of papers which are concerned with various aspects of computer application and engineering. Far too little is known of computer engineering in the Soviet Union by workers in the West and this translation of Lebedev's book will go some way to fill this gap.

The book contains eight papers dealing with the following topics: the power supply system of BESM, digital differential analysers, dynamic flip-flops, automatic checking in serial arithmetic units, word selection from a dictionary, an analysis of the use of ferrite cores in matrix storage, an investigation of the reliability of a matrix magnetic store, and, finally, some suggestions about nomenclature in digital computer engineering. The papers on power supply and basic nomenclature will not be of very much interest, particularly as the original Russian version of this book appeared in 1955. The remaining papers, however, will be of great interest to all British engineers.

The translation of the book is excellent and, although the typesetting (in photolithographed, unjustified typescript) leaves something to be desired, the book can be thoroughly recommended. A. D. BOOTH

Elektronische analogrechner. By D. ERNST. (Munchen: R. Oldenbourg, 1960.) Pp. 315. Price DM.38.

This work gives a well balanced introduction to the whole field of analogue computation. Analogue and digital methods are first compared, and analogue equipment is classified into different types. Customary forms of the basic elementary units are next discussed, with due attention to practical limitations deriving from allowable ranges of magnitude and from such matters as time constants and errors of integrators.

About half the book is devoted to programming, with six examples from electrotechnology fully worked out in practical sequence—equations, interconnection scheme, scaling factors.

Choice of equipment for different purposes and the organization of a computing establishment are handled at some length, and finally there are descriptions, illustrations and specifications of 16 instruments from six countries (including three from Great Britain).

The book is well produced, with good diagrams in ample quantity. There are 237 references, about half of which are to English-language titles. T. H. O'BEIRNE

Encyclopaedia on cathode-ray oscilloscopes and their uses. By J. F. RIDER and S. D. USLAN. (London: Chapman and Hall, 1959.) Pp. xix + 1325. Price £10 10*s*.

This is a monstrous book, in more than one sense of that adjective. It contains 1325 8½ × 11 in. pages and weighs over 7 lb. Furthermore, it is profusely illustrated, with photographs, line diagrams or graphs on nearly every page. If, therefore, it were packed with valuable information, the high price charged for the book might be justified. Such, however, is hardly the case.

Of the five sections into which the book is divided, the first (108 pages) deals with the construction and mode of operation of cathode-ray tubes. From the table of contents it appears that most aspects of the subject are discussed, but the reader soon finds that the treatment is often annoyingly superficial. For example, although five pages are devoted to practical details of magnetic-deflection coils, no mention is made of ferrite (as distinct from 'iron-dust') yokes, the cosine distribution of turns in a slotted yoke is not discussed and the reader who wishes to design a set of coils is given little useful information. Many similar examples might be given.

Criticisms of the same kind can be made of Section 2 (426 pages) which is concerned with the circuits used in oscillographs and accessory apparatus, and of Section 3 (374 pages) which is devoted to applications of oscillographs. In each section an extremely wide range of topics is touched upon in a very superficial manner. The result is likely to be completely bewildering to the novice and the details given are insufficient to form a useful work of reference for the qualified engineer or physicist.

In dealing with applications, the authors apparently adopt the standpoint that any piece of research in which an oscilloscope has been used is to be regarded as an application. From this point of view, a microwave reflectometer involving a swept-frequency klystron oscillator, directional couplers, an attenuator, crystal detectors and a ratio-meter, becomes an application of the cathode-ray oscilloscope. It is perhaps relevant to add that the description of this apparatus occupies 350 words plus a block diagram and is supposed to be intelligible to a reader who, in an earlier chapter, required ten pages of explanation concerning the way in which the spot in a cathode-ray tube is deflected electrostatically.

Section 4 (144 pages) consists of two parts. The first of these records diagrams of some 1600 waveforms which are obtained by the addition of harmonics of varying amplitude and phase to a fundamental. It is suggested that comparison of an unknown waveform with these known ones will make possible a harmonic analysis of the former. Similarly, the second half shows, by means of a large number of diagrams, the distortion produced in square waves of various frequencies when applied to amplifiers having certain stipulated frequency characteristics. It seems doubtful whether experimenters will find much use for this section.

The last section (232 pages) gives specifications and circuit diagrams of 78 commercial oscilloscopes, most of which are manufactured in the U.S.A. Once again, this information is of doubtful value.

Taking the book as a whole, it is difficult to resist the conclusion that the authors have deliberately set out to produce a volume of very large size. It cannot be said that the result justifies their efforts. C. W. OATLEY

Feedback control system analysis and synthesis. By J. J. D'AZZO and C. H. HOUPIS. (London: McGraw-Hill Publishing Company, 1960.) Pp. xii + 580. Price £5 4*s*. 6*d*.

When I reached the second chapter of this book, that on differential equations, I thought we had at last got a book

taking into account practical problems, since the governing differential equations are developed for hydraulic and thermal systems as well as mechanical and electrical. Also, the mechanical and electrical analogies are indicated, which is useful. But, alas! the remainder of the book, except for one chapter, is solely concerned with the linear theory of the design of simple control systems for determinate inputs, such as could be found in any of about twenty books already available, several from the same publishing house. This is not to say that it is not well presented and useful, but is it really necessary? Anyone who has Truxall's *Automatic Feedback Control Systems Synthesis*, published in 1955, has everything that is in this book and even if he only has Brown and Campbell's *Principles of Servomechanisms*, published in 1948, he is unlikely to learn from this book anything more which will help him in the design of control systems.

Non-linearities get short shrift, with only twenty-six pages dealing with describing functions. There is no mention of phase-plane techniques nor of any of the more modern methods of dealing with non-linear systems. In a chapter on complex control systems, the authors fail to distinguish between multi-loop and multi-variable systems and the latter are only dealt with in a very cursory manner, but it is useful that the aircraft control system, which involves interaction between the varieties, is dealt with as an example. Load variations are mentioned, but only those that occur explicitly in the governing equation of the system, which is the easy case to deal with but by no means typical. The real problems arise when the load changes are parameter changes in the transfer function of the load, which is very much more usual. Conditional feedback is mentioned but not developed, the reader being referred to an article in *Applications and Industry*, which is a rather unsatisfactory omission.

Surely the time has come when a book of this sort should put more emphasis on statistical methods of design, which are precise, rather than the 'cut and try' methods which have been the 'bread and butter' of authors in the control engineering field for so many years. Anyone would think that, in the majority of problems facing him, the control engineer knows what the input to his system is going to be, whereas it is obvious that in practice this is the exception rather than the rule, and when high performance is required it is quite essential that statistical methods be used. It is, however, accepted that in many practical cases the methods described in this book have proved satisfactory. Even so, designers faced with a practical problem usually apply these methods only as a guide, because they know they have insufficient information to make an accurate analysis and in any case they will have to make some adjustments when the system is installed. When it is installed they have access to the closed loop response and no longer need methods of investigating the closed loop response when only the open loop transfer function is known.

This book may be excellent for teaching, but those who are concerned with the design of practical systems will do better to go to the first book on the subject from the same publishing house, to wit, James, Nichols and Phillips' *Theory of Servomechanisms*, first published in 1947, which is a very much more practical book, in spite of its title.

J. F. COALES

Overvoltage research and geophysical applications. Edited by J. R. WAIT. (London: Pergamon Press, 1959.) Pp. viii + 158. Price 55*s*.

This book is concerned with the induced polarization arising when a current is passed between electrodes in the ground and its application as a method of prospecting for minerals. It is, in fact, an assemblage of ten papers, written by nine authors, one of whom is the Editor. All the contributors are, or were, associated with the Newmont Mining Corporation which has been investigating the phenomenon both in the laboratory and in the field for over ten years. It is, without doubt, the most comprehensive book on this subject and the main criticism is its lack of continuity—a feature which inevitably follows from this type of production.

After a brief history, the distribution of current in a continuous body is examined if a volume element δv of the medium acts as a current dipole of strength $-mj\delta v$ where j is the current density and m, like 'chargeability', is a constant of the medium and represents the influence of the induced polarization. Such a system will influence the observations made with d.c., a.c. and the growth and decay step functions applied to the ground. These aspects are discussed in relation to observations on the properties of rocks measured in the laboratory and applied in the field as a method of prospecting. It appears that the method has been successful particularly in the search for some disseminated metallic sulphides which yield large induced polarizations.

J. M. BRUCKSHAW

Heat

Conduction de la chaleur en régime variable. By G. RIBAUD. (Paris: Gauthier-Villars, 1960.) Pp. 90. Price 18NF.

This is a very simple and straightforward account of the mathematical and graphical method of solution of the unsteady state heat conduction equation. As far as your reviewer can see, it does not attempt to introduce new methods nor does it attempt to discuss detailed experimental work. It is purely a straightforward account of the equations and methods of their solution. The thermal properties of a number of materials are given in a table in the neighbourhood of atmospheric temperature, but the complicated problems which arise from the variability of thermal conductivity and specific heat with temperature are not dealt with in any detail. The diagrams are very clear and well drawn, and there are a number of worked out examples of iterative graphical and tabular methods of solution. One could not ask for a clearer and more accurately presented account of a limited field, and the book can be fully recommended for teaching purposes for those who wish to be sure that they will make no mistake in the solution of the equations of this type of heat transfer.

M. W. THRING

Direct conversion of heat to electricity. Edited by J. KAYE and J. A. WELSH. (New York: John Wiley and Sons; London: Chapman and Hall, 1960.) Pp. ix + 363. Price 70*s*.

This book contains a number of papers dealing with the different methods of converting heat directly to electricity. Several of the papers had already been published in 1958 and 1959, and those not previously published were presented at a meeting at the Massachusetts Institute of Technology in 1959. The Editors decided that the subject was so live that it was desirable to publish these papers quickly; no attempt has therefore been made to correlate the papers or to combine and reorganize them into a homogeneous account. There is thus a certain amount of repetition, particularly in the first two sections. The first of these deals with close-clearance vacuum thermionic diode converters, and the

second with diodes containing a low pressure of caesium which is ionized and so serves to decrease the space-charge limitation of the emission. The different points of view in these papers lead to essentially the same conclusions, although this is not obvious in all cases; it is a good exercise for the reader to study these papers and to co-ordinate them for himself.

These sections are followed by two short interesting papers about magnetohydrodynamic converters, which analyse the conditions required to make this type of generator practicable. The third section on thermoelectric converters deals with the basic theory, generator design considerations, the determination of the basic parameters of thermoelectric materials, and the detailed properties of two of these materials. A thermoelectric refrigerator is also described.

The final section surveys the developments in fuel cells as direct energy converters.

The presentation of these papers in one volume will be found useful and stimulating by the increasing number of people working on this subject. D. A. WRIGHT

Radiation pyrometry and its underlying principles of radiant heat transfer. By T. R. HARRISON. (New York: John Wiley and Sons; London: Chapman and Hall, 1960.) Pp. xi + 234. Price 96*s*.

This is a first-class specialist book primarily concerned with the statement and derivation of the mathematical equations concerned in the design and use of radiation pyrometers. It starts accordingly with three chapters discussing the laws of emission, absorption, reflection and transmission of thermal radiation, and then goes on to give in the remaining chapters mathematical analyses of the physical and operating characteristics of radiation pyrometers and the corrections for emissivity of the source. A detailed analysis is also given of the light guide radiation pyrometer using a saphire rod projecting from the furnace into a radiation instrument so that measurements can be taken where a very small hole is available up to somewhere near the melting point of alumina. Detailed tables for the calculation of the black body radiation characteristics as a function of temperature and wavelength are given, and also a set of diagrams and sample calculations. The book would therefore be invaluable to anyone concerned with the design of radiation pyrometers, or their use in special cases other than those for which they were intended. Calibration of radiation pyrometers is dealt with rather more from the theoretical aspect than from the point of view of practical details of calibration furnaces; for example, the use of carbon tubes for electric heater elements for calibrating total radiation pyrometers up to 1700° C developed in the steel industry of this country is not mentioned. M. W. THRING

Mathematical tables

Quantum chemistry integrals and tables. By J. MILLER, J. M. GERHAUSER and F. A. MATSEN. (Austin: University of Texas Press, 1959.) Pp. 1223. Price $15.

This work is an extensive collection of some integrals and tables needed in quantum mechanical calculations of the properties of atoms and molecules. It is concerned with Slater type-atomic orbitals and has been prepared to assist in obtaining any one- or two-electron, one- or two-centre 1*s*, 2*s* or 2$p\sigma$ integrals over the operators involved in energy and dipole moment calculations. The first few pages give algebraic formulae for these integrals in terms of the common auxiliary functions *A*, *B*, *G* and *W* discussed by Kotani *et al.* Numerical tables of these four auxiliary functions then follow: that for *W* is especially extensive and runs to 1125 pages. The calculations were made on an IBM 650 computer with double and triple precision floating-point interpretive routines. To eliminate errors the tables were printed by a photo-offset process from the original computer output. The tables are easy to read, with the values of the integrals given in a floating-point system. The decimal exponents however need a little getting used to. The number 12·345 which is $10^2 \times 0{\cdot}12345$, is printed as 52 12345, the exponent having been increased by 50. The authors are to be congratulated on their industry.

D. W. J. CRUICKSHANK

Nuclear physics, apparatus, techniques and data

Neutron detection. By W. D. ALLEN. (London: George Newnes, 1960.) Pp. vii + 260. Price 45*s*.

The subject of this book is one of immense practical importance, not only to the experimental nuclear physicist, but also to that growing band of reactor technologists, who are concerned with the design and operation of new reactor plants. The author has evidently found it difficult to decide to whom his book should be addressed, and has adopted a middle course, several of the earlier chapters summarizing the necessary elements of nuclear physics and particle detection. Later chapters do, however, detail those methods and techniques which the author considers important and he takes the trouble to provide an extensive bibliography and appendices.

The contents of the book may be seen from the chapter headings—Introduction; Reactions used in neutron detection; Chief instruments of neutron detection; Applications of neutron detection and neutron standards. On the whole the treatment is good so far as the experimental nuclear physicist is concerned, but the reader interested in the uses of neutron detection in nuclear reactors will find many omissions. Also the author claims that the subject-matter covers material up to the end of 1957, but there are a number of important omissions from his bibliography which would interest the reactor technologist. Nevertheless the book is a valuable addition to our Library on radiation and particle detectors, and will be found useful to both students and workers in the field alike. The standard of the printing and figures is excellent. D. TAYLOR

Reactor handbook. Vol. 1. Materials. 2nd revised edn. Edited by C. R. TIPTON. (London: Interscience Publishers, 1960.) Pp. xv + 1207. Price £13.

After five years we have a second edition of the *Reactor Handbook*. This massive volume of over 1000 pages, prepared under contract for the U.S. Atomic Energy Commission, is one of a series of four. The other three on Physics and Shielding, Engineering and Fuel Reprocessing will come out later.

The new layout now includes liquids and gases as well as solids. The scope may be seen from the section headings which are: General (safety, radiation damage, etc.); Fuel materials; Cladding and structure materials; Control materials; Moderator materials; Coolant materials and Shielding materials. The bulk of the text is devoted to fules and to cladding and structural materials. Alloys of the fissile elements and high temperature metals are discussed in some detail.

The text comprises a series of articles by specialists profusely sprinkled with graphs and tables. Each section has a long bibliography including some classified documents for the benefit of those having access to them.

Unfortunately the dates of the references only run up to early 1958 which excludes the papers presented at the 2nd Geneva Conference later that year.

It is doubtful whether many individuals will afford the £13 for this book but for library reference it must be considered as the authoritative work, particularly for metallurgists and chemical engineers.

W. S. EASTWOOD

The study of elementary particles by the photographic method. By C. F. POWELL, P. H. FOWLER and D. H. PERKINS. (Oxford: Pergamon Press, 1959.) Pp. xvi + 669. Price £12 10*s*.

During the last 10 to 15 years tremendous interest has been shown in the quest for new fundamental particles and in the evaluation of their properties. It is widely accepted that the remarkable progress made in this field of study is largely due to the exploitation of experimental techniques which reveal the tracks of charged particles. These 'visual techniques' include the Wilson cloud chamber in various forms, the photographic emulsions and several different kinds of bubble chamber. The cloud chambers and emulsions were first used with success in the cosmic-ray flux of high-energy particles. Later all the visual techniques were used in beams of particles generated by the giant accelerating machines. Nuclear disintegrations produced by particles with a wide spectrum of energies have been studied in detail, and particular attention paid to the many different mesons and hyperons produced directly or indirectly in high energy collisions.

Professor Powell and his two colleagues have themselves made many important contributions to our knowledge both of cosmic-ray and particle physics, using the photographic method. It is, therefore, most appropriate that they should write a very detailed account of the techniques used in emulsion work and also to include an authoritative account of the basic discoveries.

The scope of the book is very ambitious but, without doubt, the authors have succeeded. A detailed account of the development of the emulsion technique, from the earliest days of α-ray tracks in emulsion in 1911 to the present-day use of stacks of emulsion, is given together with a detailed discussion of the properties of the fundamental particles. The last two chapters are concerned with studies of the behaviour of primary cosmic rays in emulsions, and with nuclear collisions at energies greater than 50 Gev. At present particles of these energies are only found in reasonable numbers in the cosmic rays near the top of the earth's atmosphere. The book is copiously illustrated; the beautiful microphotographs form an atlas which illustrates both the capabilities of the different kinds of emulsion and many of the properties of the particles produced in energetic nuclear disintegrations.

This book will be of the greatest value to physicists working with the emulsion technique. It must surely be their basic reference work for many years to come, both for the history of the subject and for the detailed discussion of the techniques as currently employed. The references are numerous and well chosen. The book, however, is not solely for the expert, or would-be expert. It can be strongly recommended to all workers in high-energy nuclear physics, particularly those who now use bubble chambers.

The production is excellent and there appear to be few errors, a remarkable achievement in such a large and detailed work.

C. C. BUTLER

Pure and applied physics, Vol. 9. Nuclear spectroscopy. Parts A and B. Edited by F. AJZENBERG-SELOVE. (London and New York: Academic Press, 1960.) Part A, pp. xxi + 621. Part B, pp. xv + 526. Price £5 14*s*. 6*d*. each.

The detailed study of energy levels of nuclear systems has now been going on intensively for some time. The volume and complexity of the information now available is very great and must indeed be terrifying to the newcomer in the field. Many attempts have been made to present this information in predigested form in periodical reviews or as card indexes. The editor of the book under review has herself done yeoman work in a series of review articles. The present book tries fairly successfully to put all this in clear and handy form. It can only be partially successful because there is as yet no unified theory to put in the first chapter to which data could be related. In fact the theories (four of them) are relegated to the last section.

The first volume has chapters by twenty-two of the most experienced experimentalists on all the techniques which have been used to find things out about energy levels. The main parameters to be determined are energy, angular momentum, parity, isotopic spin, electric and magnetic moments and the lifetime of the excited state. In describing the experimental methods, the authors in fact survey practically the whole field of low energy nuclear physics.

In the second volume, nine authors show how the experimental data can be analysed to give more basic information and correlations. Here, theoretical guesses about the structure of the nucleus or about the nature of nuclear matter begin to intrude. Finally the nuclear models which attempt to correlate large areas of the experimental data are described —the shell model, intermediate coupling model, collective model and complex potential model.

One can hope that by the time another edition of this book is prepared the number of models required will be reduced to one. In that case, however, nuclear spectroscopy would be much less interesting. In the meantime the present book will be invaluable to the working nuclear physicist.

E. B. PAUL

Experimental nuclear physics. Vol. 3. Edited by E. SEGRÈ, G. C. HANNA, M. DEUTSCH, O. KOFOED-HANSEN and E. M. MCMILLAN. (New York: John Wiley and Sons; London: Chapman and Hall, 1959.) Pp. ix + 811. Price £9 4*s*.

This is the third and final volume of a useful advanced text book on practical nuclear physics. Two-thirds of the present volume is devoted to a detailed survey of the phenomena of radioactive decay. The approach is that of the experimentalist. Sufficient theory is included to enable the experimental results to be interpreted. This portion is introduced by E. Segrè in a chapter summarizing the mathematical formalism used to describe radioactive decay. Alpha radioactivity is the subject of another chapter by G. C. Hanna which is a model of completeness and clarity. Martin Deutsch and Otto Kofoed-Hansen have jointly prepared the next two sections on gamma rays and beta rays. Although gamma-ray decay resulting from a nuclear reaction is an

important phenomenon, the authors have chosen to concentrate on gamma ray phenomena resulting from radioactive decay. Though the theory is, of course, the same no matter how the nucleus became excited, the experimental material is somewhat different. The chapter on beta rays is a clear and useful exposition of the whole subject. Many new experimental techniques were developed in the period of feverish activity following the discovery of the violation of parity in β-decay. This is treated in an extensive section on experiments concerning neutrinos.

The final third of the book is an article on particle accelerators by E. M. McMillan. This differs considerably in approach from the rest of the book. A large part is devoted to the early history of accelerators and much of the remainder is descriptive material. Much of this is very interesting, but, for only a few machines does it give enough details to understand why they work. Although in the preface it is stated that only nuclear phenomena below the meson threshold are included, in the accelerator chapter fairly long accounts of machines in the Gev range are included. E. B. PAUL

Atomkraft. By F. MÜNZINGER. 3rd edn. (Berlin: Springer-Verlag, 1960.) Pp. xii + 304. Price DM. 42.

The author, a well-known German power-station engineer, calls his book a critical introduction for engineers, economists and politicians. From this aspect it can be considered a valuable contribution to the existing literature. It is not written for workers in the field, and refinement in the theoretical presentation has been displaced by the desire for simplicity. Mathematical treatment has been avoided whenever it seemed possible. The economic aspects are rather strongly emphasized. A large amount of information has been collected and references are given for more detailed studies.

The first chapter is an introduction to fundamental principles of nuclear energy and an attempt is made to give a simple qualitative presentation of the different phenomena and requirements. The second technical chapter is a rather comprehensive survey of the different types of reactors including a description of representative projects. The principal difficulties and the limitations of the different approaches are stated and associated problems like waste disposal and possible hazards are briefly dealt with. A short article on the possibility of fusion is also included. In the third chapter the economics of nuclear power are discussed in detail, mostly based on American publications. The situations in the different countries are compared and some extrapolations for the prospects in West Germany are made. A comparison of the various national atomic power programmes is given. The second and third chapters are particularly useful references for general information.

The last chapter deals briefly with the aspects of mobile atomic power units. H. SCHMIDT

Nuclear photo-disintegration. By J. S. LEVINGER. (London: Oxford University Press, 1960.) Pp. 144. Price 15*s*.

Several excellent review articles covering the field of photo-nuclear reactions have appeared in the literature in recent years. The present work is, however, the first book devoted to this subject, and in it Professor Levinger has attempted not simply another review of the theoretical or experimental material in the literature, but rather a logical development of the theory of nuclear photo-disintegration.

The development starts with calculations of the atomic photo-effect and thence to the nuclear effect for cases where the analogy is valid. This is followed by chapters discussing the photo-disintegration of the deuteron, sum-rule calculations, discrete transitions, total photon absorption cross sections, and the products of nuclear photo-disintegration. The applicability of various nuclear models in describing the observed phenomena is discussed. The book contains only brief mention of experimental techniques, experimental results being quoted to illustrate the validity of the calculations. The bibliography contains 224 references up to 1959.

There can be no doubt that this excellent little book, which is published at a popular price, will find its way as a reference work on to the bookshelf of any physicist who is interested in the interaction of electromagnetic radiation with atomic nuclei. E. R. RAE

Optics

The technique of photomicrography. By D. F. LAWSON. (London: George Newnes, 1960.) Pp. xvi + 256. Price 55*s*.

This is a useful laboratory manual. It is entirely non-mathematical, and concentrates upon providing information for use at the bench. All parts of the essential optical equipment are described, including (as is characteristic of books of this kind) much that could equally well be found in manufacturers' catalogues. Judged by present-day standards, the illustrations, especially the colour reproductions, seem rather poor; this applies more especially to the plates, the line-drawings are much better. There are some curious vagaries; for instance the treatment of phase-contrast and of interference methods is excellent, whereas that of Conoscopic figures in crystals is most obscure. The latter not only says little about what is happening, but fails to bring out the special arrangement of condenser and lens units needed for success.

Special methods of illumination, like infra-red and fluorescence techniques are discussed, with data relating to filters, transmission limits and so forth to correspond. At the end is a valuable glossary of terms, classified bibliography, and an adequate index. F. I. G. RAWLINS

Space and aircraft technology

Space technology. Edited by H. S. SEIFERT. (New York: John Wiley and Sons; London: Chapman and Hall, 1959.) Pp. xvi + 1000. Price £9.

This book contains a large amount of useful material for anyone working on any aspects of space exploration. It includes so much that it could not fail to contain a great deal of interest for any scientist or engineer even though not a specialist in the field. The techniques required in space exploration are so wide that many branches of astronomy, aerodynamics, physics, chemistry and biology are involved.

The material included in the book covers the content of a comprehensive lecture course given in the University of California. It is divided into five main sections, the first dealing with flight dynamics, the second with propulsion and vehicle structure engineering, the third with communications and guidance, the fourth with environmental problems of man in space and the last with scientific and other applications. There is even included in the appendix a list of one hundred questions on the subject-matter of the course.

Two introductory chapters discuss the reasons why the

subject of space exploration is of interest and importance. Nothing especially new appears but it is nevertheless interesting to see expressed in print some of the reasons why there is such great activity in the subject in the United States.

The section on flight dynamics includes chapters on flight performance of a rocket, trajectory optimization, earth satellites and perturbation theory, the vanguard satellite trajectories and orbits, lunar flight trajectories (a very useful chapter), interplanetary observatories, low-thrust flight, time dilation effects in space travel, heat transfer at hypersonic speeds in a planetary atmosphere, and the possibility of a safe landing. This is followed in the next section by chapters on chemical rocket fundamentals, liquid propellant engines, solid rocket propulsion, nuclear rocket propulsion, magnetohydrodynamics, structural configurations and materials and integrated design analysis. The next section contains chapters on the feasibility of space communication, space communication implementation problems, problems of radio guidance, and inertial, radio-inertial and midcourse and terminal guidance. Chapters on the physical factors of the space environment, medical aspects of manned space flight, space cabin design and crew performance comprise the next section. In the final section there are chapters on the I.G.Y., research in the upper atmosphere, scientific uses of space vehicles and what the future holds. The latter is a compendium of views expressed by members of a panel, each of whom is closely connected with one or other aspect of the subject.

It is inevitable that the different chapters are somewhat uneven, but there is an immense amount in the book which will be invaluable as a reference text to which one may turn for information on almost any branch of a very broad subject. The value is enhanced by the provision of a list of references after each chapter. Even those who, for one reason or another, regard space research as an over-expensive luxury might find the book of much interest as a text in applied science.

H. S. W. MASSEY

Leichtbau. Bauelemente, Bemessungen und Konstruktionen von Flugzeugen und anderen Leichtbauwerken. By H. HERTEL. (Berlin: Springer-Verlag, 1960.) Pp. xxviii + 526. Price DM. 67.50.

Teaching and research in the field of aircraft technology were interrupted in Germany for a period of about 12 years following the end of the Second World War, and when they were resumed at the end of 1955 the need for a modern textbook was apparent. This book was written to fill the need and is based on the author's lectures given at the Technical University, Berlin. It is a book written by a designer for designers, and mathematical derivations and theories are not included. Nevertheless, theoretical knowledge and practical experience are dealt with together in order to clarify principles and to provide a basis for new design. Exhaustive treatments of basic constructional problems should give the student designer a clear understanding of fundamentals and enable him to tackle future problems successfully.

In spite of a comprehensive treatment of conventional aero-engineering practice the book does not expound new problems arising in this rapidly advancing industry. For example, no reference could be found to aero-dynamic heating in high-speed flight. The book has a very detailed and systematically subdivided list of contents running to 20 pages, but the absence of an index will severely handicap the user.

W. BETTERIDGE

Symposia and conference proceedings, etc.

Viscoelasticity: phenomenological aspects. Edited by J. T. BERGEN. (New York and London: Academic Press, 1960.) Pp. x + 150. Price 48*s*.

This book consists of the seven papers presented at a symposium on the phenomenological aspects of viscoelasticity which was held on 28th and 29th April 1958 at the Research and Development Centre of the Armstrong Cork Company, Lancaster, Pennyslvania.

Solid state physics in electronics and telecommunications. Vols. 1, 2 and 3. Edited by M. DÉSIRANT and J. L. MICHIELS. (London and New York: Academic Press, 1960.) **Vol. 1, Semiconductors, Part 1:** Pp. xxiii + 638. **Vol. 2. Semiconductors, Part 2.** Pp. xvii + 646. **Vol. 3. Magnetic and optical properties, Part 1.** Pp. xvi + 557. Price £6 8*s*. 6*d*. each volume.

These books consist of the published proceedings of an international conference held in Brussels from 2nd–7th June 1958. There were some 800 active participants at this conference coming from twenty-three countries and 250 papers were presented.

The papers in the first volume are on the preparation and properties of semiconductors, solid state theory, effects of intense electric fields, noise and surface phenomena. Volume 2 is on semiconductor compounds, applications, devices and power rectifiers. Volume 3 is on ferrites, properties of magnetic materials, garnets and magnetic resonance and the fourth volume will contain the remaining papers on magnetic and optical properties.

Plasma dynamics. Edited by F. H. CLAUSER. (London: Pergamon Press, 1960.) Pp. ix + 369. Price 84*s*.

A symposium on plasma dynamics, sponsored by the U.S. Air Force Office of Scientific Research and prepared under the auspices of the U.S. National Academy of Sciences, was held in Massachusetts in June 1958. There are nine chapters under the following titles: Astrophysics; Fluid mechanics; Thermonuclear physics; Gaseous discharges; Electron beam dynamics; Statistical mechanics.

Structural damping. Edited by J. E. RUZICKA. (Oxford: Pergamon Press, 1960.) Pp. iv + 165. Price 30*s*.

This book contains the papers presented at a colloquium on structural damping, sponsored by the Shock and Vibration Committee of the Applied Mechanics Division of the American Society of Mechanical Engineers, was held at the ASME annual meeting in Atlantic City, in December 1959.

Photochemistry in the liquid and solid states. Edited by L. J. HEIDT, R. S. LIVINGSTON, E. RABINOWITCH and F. DANIELS. (New York: John Wiley and Sons; London: Chapman and Hall, 1960.) Pp. vi + 174. Price 40*s*.

This book is based on a symposium sponsored by the National Academy of Sciences and the National Research Council. The authors discuss photochemistry storage, photochemical reactions, and the requirements for reaction types which might prove useful for storing solar energy.

Medical electronics. Edited by C. N. SMYTH. (London: Iliffe and Sons, 1960.) Pp. xxiii + 614. Price £7 5*s*.

The proceedings of the Second International Conference on

Medical Electronics, held in Paris from 24th–27th June 1959, are contained in this book. For convenience in publication the papers have been re-grouped into eight broad sections under the following headings: Electrophysiological techniques; Electroencephalography; Cardiology; Manometry and flow measurement; Acoustic techniques; Automation in medicine; Radiology and isotopes; Chemical instrumentation.

Immediate and low level effects of ionizing radiations. Edited by A. A. BUZZATI-TRAVERSO. (London: Taylor and Francis, 1960.) Pp. xii + 381. Price £2 15*s*.

Under the joint sponsorship of UNESCO, IAEA and CNRN a symposium was held in Venice in June 1959 on the immediate and low level effects of ionizing radiations. The proceedings have now been published as a supplement to the *International Journal of Radiation Biology* containing the thirty-three papers appearing in their order of presentation at the meeting.

Electromagnetic wave propagation. Edited by M. DÉSIRANT and J. L. MICHIELS. (London and New York: Academic Press, 1960.) Pp. xiii + 730. Price £7 17*s*.

This volume contains the papers presented at the international conference sponsored by the Postal and Telecommunications Group of the Brussels Universal Exhibition held in 1958. Over fifty papers are presented by authors from eleven different countries.

Gas chromatography abstracts. Edited by C. E. H. KNAPMAN. (London: Butterworths Scientific Publications, 1960.) Pp. ix + 164. Price 42*s*.

This volume of abstracts is the first to be compiled by the Abstracting Service organized by the Gas Chromatography Abstracts Editorial Committee of the Gas Chromatography Discussion Group, and it is a continuation of the work already published in the inaugural volume *Gas chromatography abstracts* 1958.

Theoretical physics

The statistical theory of phase transitions. By B. T. GEILIKMAN. Translated from the Russian. (New York: Consultants Bureau; London: Chapman and Hall, 1959.) Pp. 77. Price 32*s*.

This is a book for theoretical specialists only, and indeed only for those theoretical specialists who can cope happily with irritatingly bad, or singularly infelicitous, English and typography. The reviewer, alas, has not proved fully capable of so doing: though perhaps he is asking for a freer translation and higher standard of production than are practicably feasible. Suffice it to say that the work appears to be a serious attempt to reckon with the volume-dependence of the cluster integrals in Ursell–Mayer imperfect gas theory. The book was published in Russia in 1954. Since then an independent attack on this problem has been made by Ikeda (*Progress of Theoretical Physics*, 1956 and later), and anyone wishing seriously to enter this difficult field should certainly compare the two approaches. It is a pity that the work under review is as agressive as it is in tone; for the pioneer workers, Mayer, Kahn and Uhlenbeck, Born and Fuchs, particularly Mayer, were not only breaking new ground but setting a quite new standard of achievement in statistical theory. Nevertheless all translations of significant Russian work are to be welcomed. The present translation was produced by Consultants Bureau, New York, in 1957, but has only recently been published at the present reasonable price. The reviewer regrets that there is no indication in it of other monographs obtainable in the same series.

G. S. RUSHBROOKE

Note. We have also received copies of several new journals, details of which are given under *Notes and Comments* on page 83.

Investigation of periodic discontinuities in the mutual potential energy of molecules of water and some polyhydric alcohols

by M. M. QURASHI and A. K. M. AHSANULLAH, Physics Division, Central Laboratories of the Council of Scientific and Industrial Research, Karachi, Pakistan

MS. received 9th February 1960, *in final form 2nd November* 1960

Abstract

The inter-molecular energy of activation of viscous flow has been measured at close intervals of 2 to 5 deg C for glycerol, ethylene glycol and thrice-distilled water, using the logarithmic derivative of the Andrade (1934) *equation, namely,*

$$\frac{\Delta \ln \eta}{\Delta T} = -\frac{\epsilon/k}{T^2}.$$

In the case of water and ethylene glycol, the temperature has been controlled to 0·002 and 0·005 deg C respectively, and the temperature intervals measured with a calibrated Beckmann thermometer. By maintaining a corresponding accuracy in the viscosity measurement, the values of ϵ/k for water are reproducible to 0·1% in the temperature range from 6 to 45° C and to 0·3% from 45 to 72° C. The results when plotted against temperature indicate a step-like variation of the energy ϵ, the repetition interval between steps being 25 deg C for glycerol, 12 deg C for ethylene glycol, and 6·5 ± 0·2 deg C for water.

The drop between successive steps in water is sharp, extending only slightly more than the measuring interval of 2 deg C, and each step is flat within the limits of experimental error. The energy interval $\Delta\epsilon$ between steps is not constant, but decreases steadily towards higher temperatures. The average value of this interval $\Delta\epsilon$ is of the order of 80 k, i.e. 0·007 ev, and is much smaller than the thermal energy kT, but is of the same order as the fine structure observed in the OH absorption of water at 3300 cm^{-1}.

The temperature variation found for ϵ/k is compared with the graph calculated from the viscosity data of Bingham and Jackson (1918–19) *on water, and fair agreement is obtained when the 2 deg C measuring interval of the present experiments is smeared out to 6 deg C.*

1. Introduction

THE ultimate seat of the viscous forces in all accepted theories of viscosity of liquids is taken to be the interaction energy ϵ between neighbouring molecules or molecular aggregates, which then enters through the Boltzmann distribution factor to give a formula of the type

$$\eta = A \exp(\epsilon/kT) \qquad (1)$$

where A has the dimensions of Planck's constant/volume, and is (apparently) proportional to the concentration of the interacting groups. On the earlier theory of Andrade (1934), ϵ is simply the mutual potential energy at contact, while on the more sophisticated theories of Eyring and Lennard-Jones, it measures the height of the potential barrier that has to be overcome for the relative movement of one layer with respect to the next, through a distance equal to the diameter of one molecular aggregate (Glasstone *et al.* 1941, Taylor and Glasstone 1951), and for most liquids will be a function of the temperature. For a fuller understanding of the mechanism of the viscous interaction in these liquids, it is therefore desirable to make an accurate investigation of the dependence of the energy ϵ on the molecular structure and temperature for several pure liquids, and on the temperature and concentration in the case of solutions. Preliminary experiments on two different grades of mineral oils (Qurashi 1958) had indicated a periodicity in the temperature variation of ϵ and A, and a detailed study of this feature has since been made through more accurate measurements with the polyhydric alcohols, glycerol and ethylene glycol (Naseem and Qurashi 1958, Rauf and Qurashi 1959), in which there are indications of a steep drop in ϵ from one segment to the next.

The measurements are based on Eqn (1), the logarithmic form being

$$\ln \eta = \ln A + (\epsilon/kT)$$

from which it is usual to obtain ϵ/k as the slope of a graph for $\ln \eta$ against $1/T$, but this graphical method is not adequate for studying the small-scale variations of ϵ, and analytical methods have therefore been used. Differentiation of the above equation with respect to T, in a range where ϵ and A are sensibly constant, yields

$$\epsilon/k = -T^2(d/dT)(\ln \eta), \qquad (2)$$

which gives for small finite temperature intervals ΔT,

$$\epsilon/k = -T^2(\Delta \ln \eta/\Delta T). \qquad (3)$$

For an ideal monatomic liquid, as also for other liquids in regions where both A and ϵ are constant, Eqn (3) is exact, while, in all other cases, this equation can be taken to give the mean value of ϵ/k in the temperature interval because it is equivalent to calculating the slope of the chord of the curve for $\ln \eta$ against $1/T$, instead of the tangent.

The experimental results for glycerol and ethylene glycol are shown in Figs 1(*a*) and (*b*), and the horizontal portions of the graphs indicate that these hydroxylic liquids have constant ϵ (and constant A) over certain short, regularly recurring ranges of temperature, and therefore simulate an ideal non-associated liquid in these ranges. Neither this feature nor the phenomena of recurring segments have been recognized by other workers. Since these appear to be of importance in the theory of the structure of liquids, it was considered desirable to extend the measurements to water, and also to attempt a correlation of the present type of measurement with accurate viscosity data obtained by earlier

investigators on water, which is the only simple liquid whose viscosity has previously been measured with sufficient accuracy over a wide temperature range. These experiments, which

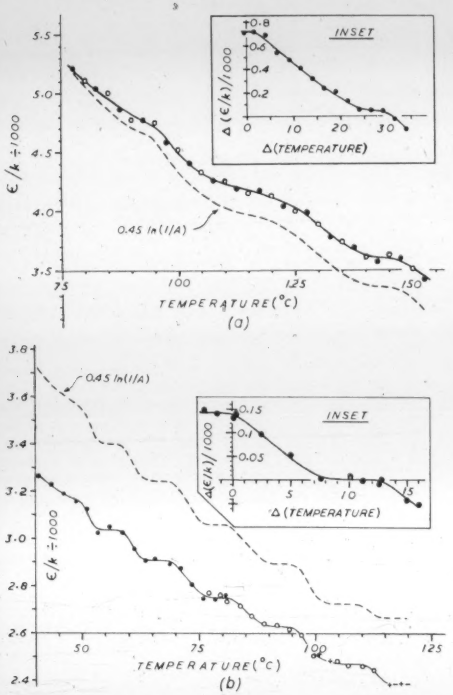


Fig. 1. Plots showing the temperature dependence of the activation energy ϵ of viscous flow for (*a*) glycerol and (*b*) ethylene glycol. The broken line curves show the parallel variation of $0{\cdot}45 \ln(1/A)$ with temperature. Insets show the averaged variation over one segment.

are described below, were expected to be particularly interesting because of a probable connection with the molecular association and other anomalous properties exhibited by water.

2. Experimental technique

Equation (3) requires the accurate measurement of $\Delta \ln \eta$, i.e. $\ln(\eta_2/\eta_1)$, where η_1 and η_2 are the viscosities at the two temperatures differing by ΔT. For the present experiments, it is easier and more accurate to measure the kinematic viscosity $\nu = \eta/\rho$ and to calculate $T^2\Delta \ln \eta/\Delta T$ by means of the formula,

$$T^2\frac{\Delta \ln \eta}{\Delta T} = T^2\frac{\Delta \ln \nu}{\Delta T} + \frac{T^2}{\rho}\frac{\Delta \rho}{\Delta T} = -\epsilon'/k - \beta T^2 \qquad (4)$$

where β is the coefficient of dilatation of the liquid and forms a small correction term. The measurements of kinematic viscosity ν were performed with an ordinary U-tube viscometer of the British Standard Specifications pattern with precautions against contamination of the test liquid, etc. The corrections for changing volume of the liquid and other factors have been discussed in previous communications (Qurashi 1958, Naseem and Qurashi 1958). When the temperature is controlled to 0·002 deg C, it is possible to attain an accuracy of 1 part in 20 000 in the viscosity measurements, thus giving an accuracy of better than 0·005 in the experimental value of $(\epsilon/k)/1000$, which is of the order of 1·8 for water at room temperature. The temperature intervals ΔT were measured with a Beckmann (differential) thermometer, the six-degree scale of which was calibrated at intervals of 0·5 deg C by comparison with a standard 1·0 deg C interval. This differential calibration curve, obtained as the mean of two experiments (solid circles and hollow circles) is plotted in Fig. 2, where the ordinates show the error in one scale

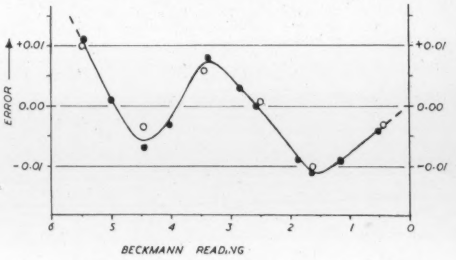


Fig. 2. Differential calibration curve for the Beckmann thermometer showing the error in the value of a Beckmann degree at different parts of the six-degree scale.

degree as measured at different parts of the Beckmann scale. The deviations of the points from the curve show that the curve is reliable to somewhat better than 0·002 deg C.

The U-tube viscometer was supported in a Townson and Mercer thermostat, with which it was found possible to control the temperature to such a degree that the Beckmann thermometer never showed a variation of more than 0·002 deg C, the mean temperature calculated from ten or more Beckmann readings for each measurement of flow time being then obtained correct to one half of this. The suction of the water into the capillary limb of the viscometer was controlled with a ballast bottle (A in Fig. 4), and the viscosity measurements were taken with rising and falling temperature sequences (Naseem and Qurashi 1958, Rauf and Qurashi 1959) each flow time being measured thrice at temperatures below 30° C and five times at higher temperatures. Each series of observations was, in general, repeated with two different samples of liquid, both to confirm the reproducibility of the results and to obtain an experimental estimate of the standard deviation of the measurement of ϵ/k. Two series of experiments have been carried out in the case of water, one with twice-distilled conductivity water and one with the same water distilled three times in order to determine any variation with repeated distillation.

3. Preliminary experimental results with water

The results of the first series of experiments with twice-distilled water, corrected for level variation, but uncorrected for Beckmann calibration, are shown in Fig. 3 (open circles), successive curves being displaced upwards through 0·1 unit. These measurements were made with viscometer No. 0 (constant = 0·000 734), with a measuring interval

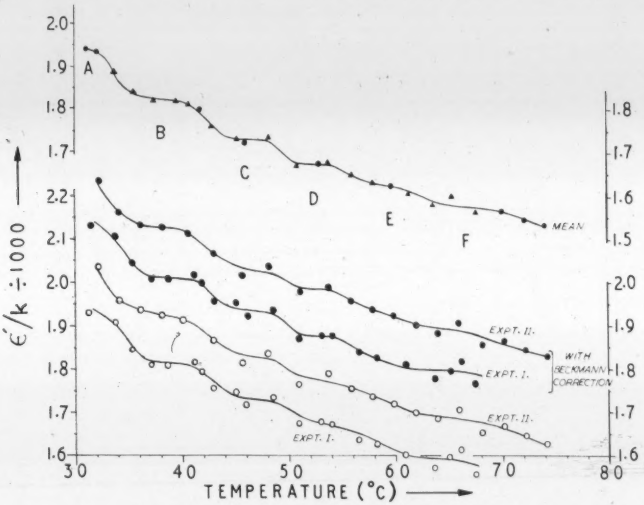


Fig. 3. Graphs of $(\epsilon/k)/1000$ obtained from the preliminary experiments on water without control of evaporation.

○ uncorrected values.
● after application of Beckmann calibration correction; successive curves are displaced upwards by 0·1 units.
△ mean of the corrected values from the two experiments.

$\Delta T = 2$ deg C. The lower part of Fig. 3 also shows the best curves drawn through the experimental points after application of the calibration correction for the Beckmann scale (solid circles). These corrected curves, as well as the standard deviations of the points, are essentially the same as the previous uncorrected ones, because the calibration correction produces a significant effect only in the third decimal place. The means of the corrected sets are plotted as solid triangles and the deviations of the points from the mean curve drawn through them in the figure are all less than 0·02, the standard deviation being nearly 0·009.

All the five graphs exhibit the presence of identical steps, the interval between successive steps being very nearly 6 deg C. From the mean graph for the two corrected sets of readings, drawn in the upper part of Fig. 3, with the observed steps marked as A, B . . . F . . ., it is apparent that, although the shape of the steps closely resembles that for ethylene glycol in Fig. 1(*b*), the energy interval between successive steps is not constant, but decreases with rising temperature, while the temperature interval remains nearly constant. In order to study these effects in more detail, it was necessary to take measurements at closer intervals and to further increase the accuracy of the viscosity readings by controlling the evaporation of the water in the viscometer, which in spite of the applied correction appears to be the largest source of error, especially at the higher temperatures. It was found possible to do this with the simple device shown schematically in Fig. 4.

4. Accurate measurements (with control of evaporation)

A small (250 cm^3) bottle B filled to about half its height with distilled water was connected on the open side of the viscometer as shown in Fig. 4. By adjusting the depth of immersion of this bottle in the thermostatic bath, it is possible to arrange approximate vapour equilibrium inside the viscometer, thus preventing loss of water from it and virtually making it a closed system. Except when sucking-up the water in the capillary limb for taking viscosity measurements, the tube C leading from bottle B to the atmosphere was kept plugged with a glass stopper.

The measurements were carried out on thrice-distilled water at temperatures from 6 up to 72° C, and were divided into two ranges, viz. from 30 to 72° C and 30 down to 6° C, (viscometer No. 0, constant = 0·000734), ice-cold water being circulated by a pump through the cooling chamber of the thermostat when working in the lower range. Experimental data for Beckmann readings, together with flow times and the values of $(\epsilon/k)/1000$, corrected only for level variation and error of the Beckmann scale, are given in Table 1 along with their standard errors, while Table 2 gives the mean values of $(\epsilon/k)/1000$, together with the errors as estimated from the differences between readings taken during rising temperature and falling temperature sequences, the standard deviations being calculated for groups of ten measurements. The smoothed corrections for expansion of the water (cf. Eqn 4) are given at the bottom of Table 2 for some temperatures, but have not been applied because it is not certain that they are always accurate in the third decimal place. The rate of increase of this correction per 10 deg C is seen to be nearly $0{\cdot}0104 \pm 0{\cdot}0005$. It therefore appears that density measurements of a much higher order of accuracy than those available are needed to bring out the existence of any sharp changes in the values of $\beta = -\partial \ln \rho / \partial T$.

Table 1 indicates values of 0·002 to 0·005 for the standard

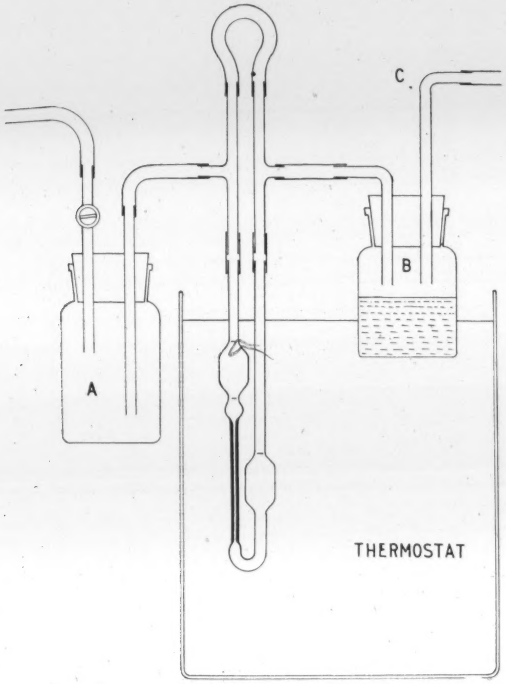


Fig. 4. Arrangement for controlling evaporation of the water from the viscometer; A is the air ballast bottle and B is half-filled with the sample of conductivity water under test in the viscometer. Variation of the depth of immersion of B gives adequate control of the evaporation from the viscometer.

deviation of the overall means of $(\epsilon/k)/1000$, and Table 2 shows that, for temperatures below 45° C, the standard deviation is now nearly constant at 0·0024, corresponding to an accuracy to very nearly 1 part in 1000 in the measurement of the quantity ϵ/k. Above 45° C, the standard deviation in Table 2 increases steadily, becoming as high as 0·007 at 70° C, which is nevertheless much better than that attained previously, and is acceptable for the present investigation.

5. Discussion of the experimental data

The measurements are plotted in the lower part of Fig. 5 (solid circles) together with the mean graph of Fig. 3 (curve only), which is displaced downwards by 0·1 unit. The agreement between the two graphs is satisfactory, and the deviations of the points about the new graph are only a fraction of the previous ones. This improvement is further reflected in the fact that the horizontal portions of the steps are now clearer and the drop from one step to the next is sharp, mostly extending over just 2 deg C. The values of $(\epsilon'/k)/1000$ and of the temperature at the middle of these flat portions are collected in Table 3, in which the unusually close steps at 11·2 and 14·3° C have been labelled as 2(*a*) and 2(*b*). (The precise value of the temperature is best determined by producing linearly the steep parts of the steps to meet the horizontal regions.) The values of $(\epsilon/k)/1000$ vary from 1·6 to 2·4, which are equivalent to an activation energy of 3 to 5 kcal/g mol ($\sim 0{\cdot}2$ ev), and would therefore correspond to the breaking of one hydrogen bond between neighbouring molecular aggregates.

In the inset to Fig. 5 are plotted the energy and temperature at the middle of the flat (solid circles and hollow circles, respectively) against the ordinal number of the step. The energy interval is seen to decrease steadily with increasing temperature (solid circles), while a fairly good straight line passing through the origin is obtained for the temperature interval (hollow circles). This straight line has a slope of $6{\cdot}5 \pm 0{\cdot}2$ deg C per step, comparison of which with the corresponding figures for ethylene glycol (12 deg C per step) and glycerol (25 deg C per step) brings out the significant fact that the three values are in the ratio of 1 : 2 : 4. The averaged shape of one step in the temperature ranges of 6 to 35° C and 35 to 70° C is shown in Fig. 6, and comparison with the insets to Figs 1(*a*) and (*b*) brings out the relative sharpness of the drop between steps, which here extends over only $3 \pm 0{\cdot}5$ deg C, i.e. 1 deg C greater than the measuring interval. The energy ϵ is usually believed to decrease uniformly and smoothly with increasing temperature

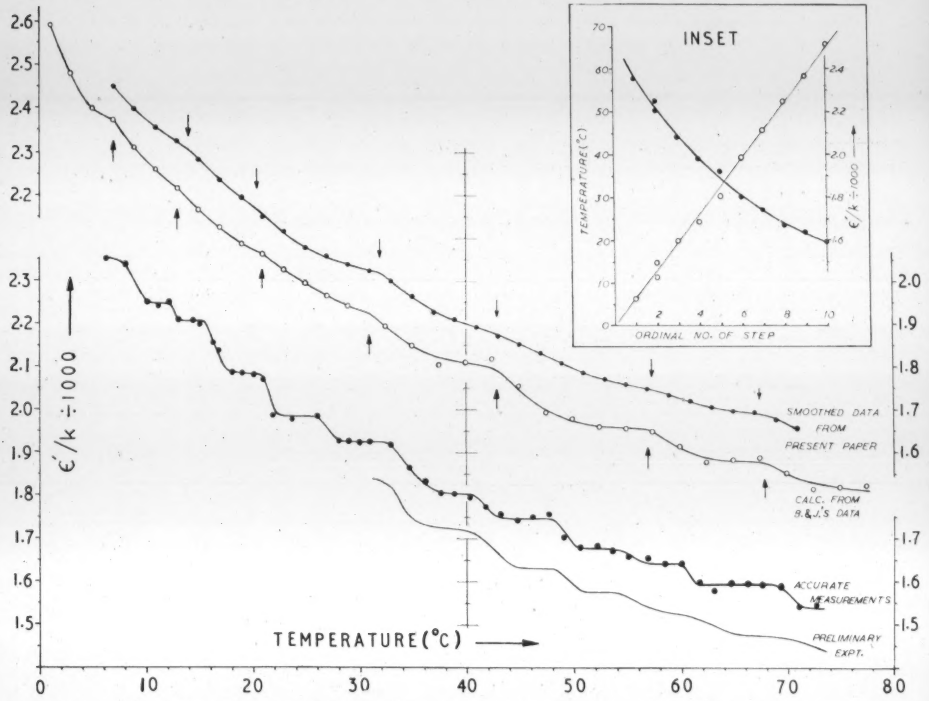


Fig. 5. *Bottom.* Plotsof $(\epsilon'/k)/1000$ obtained from the more accurate measurements on thrice-distilled conductivity water compared with the mean graph of Fig. 3 (preliminary experiment), which is shown displaced downward by 0·1 unit.
Top. The lower curve ○ is deduced from the standard data of Bingham and Jackson, and the other ● is calculated from the present experiments taking $\Delta T = 6$ deg C; the arrows mark the limits of each segment.
Inset. Plots of the value of $(\epsilon'/k)/1000$ at each step (●) and the temperature at the middle of the step (○) against ordinal number of the step.

for associated liquids, but the present investigations show that, at least for water, the energy–temperature curve is made up of a series of sharp, regularly spaced steps. The viscous behaviour over each flat (about 5·5 deg C in extent) is accurately represented by the Andrade (or Eyring) exponential equation (1) with constant coefficients A and ϵ, and this fact, while confirming the utility of this relation for the case of associated liquids, also opens up the field for interpretation of the observed drop between successive steps. On Eyring's theory, ϵ = free energy of activation and $A = (h\lambda_1/\lambda) \div (\lambda_2\lambda_3)$, where λ_1, λ_2, λ_3 are parameters that define the linear extension of the molecular aggregates, and λ is the distance between two equilibrium positions in the direction of motion.

Experimentally, one finds for the case of water that the average value of the drop $\Delta\epsilon$ between steps is 0·08 units of $(\epsilon/k)/1000$, i.e. 0·007 ev per step, which is of the order of the fine structure observed in the OH absorption of water at 3300 cm^{-1}, but is considerably lower than the thermal energy kT, and the energy of 0·025 ev corresponding to the innermost translational Raman line (Cross *et al.* 1937) (200 cm^{-1}). It is also lower than the energy transfers so far observed in neutron diffraction measurements (Brockhouse 1957) with water. Also, since the temperature interval ΔT per step is $6{\cdot}5 \pm 0{\cdot}2$ deg C, it follows that $\Delta\epsilon/k$ (which is nearly 80 per step) is of the order of 10 ΔT, i.e. $\Delta\epsilon \simeq 10\,k\Delta T$. The corresponding value of the drop in $\ln(1/A)$ can be calculated to be (Rauf and Qurashi, 1959)

$$\Delta \ln(1/A) = \Delta(\epsilon/k)/\bar{T} = 80/313 = 0{\cdot}26 \qquad (5)$$

whence the mean ratio A_1/A_2 for one step works out to be 1·29, which does not appear to be a simple fraction, but may approximate to 4/3 or 5/4, which would be consistent with jumps between discrete values of λ_1, λ_2, λ_3.

The experimental results (cf. Figs 1(*a*) and (*b*)) have already shown that the quantities $\ln A$ and ϵ/k are closely related, $\ln 1/A$ being approximately equal to constant plus $2{\cdot}5\,(\epsilon/k) \times 10^{-3}$, and their variations parallel each other. One can thus visualize the steps as being primarily caused either by a level structure in the energy ϵ, or by the existence of discrete values of $\ln A$, through the parameters λ_1, λ_2, λ_3 which define the linear dimensions of the molecular aggregates. The latter approach, i.e. discrete sizes of the aggregates, is probably more appealing, but, in view of the

Table 1. *Beckmann readings, flow times and the calculated values of* $(\epsilon'/k)/1000 = -T^2(\Delta \ln v/\Delta T)/1000$ *for thrice-distilled water at various temperatures from 5 to 73° C*

Heating sequence					*Cooling sequence*				
Temperature (° C)	*Beckmann reading (corrected)*	*Time of flow (s) corrected*	*Mean temperature (° C)*	*$(\epsilon'/k)/1000$ (not corrected for density variation)*	*Temperature (° C)*	*Beckmann reading (corrected)*	*Time of flow (s) corrected*	*Mean temperature (° C)*	*$(\epsilon'/k)/1000$ (not corrected for density variation)*
5·15	4·152 ± 0·000	2057·94 ± 0·03	6·22	2·352 ± 0·0024	5·15	4·152 ± 0·001	2057·86 ± 0·05	6·22	2·348 ± 0·0022
7·30	1·952 ± 0·001	1925·84 ± 0·08	8·12	2·334 ± 0·0024	7·30	1·952 ± 0·001	1925·96 ± 0·03	8·12	2·345 ± 0·0022
8·95	0·314 ± 0·001	1834·95 ± 0·05			8·95	0·314 ± 0·001	1834·60 ± 0·12		
9·20	5·075 ± 0·001	1816·80 ± 0·15	10·20	2·252 ± 0·0040	9·20	5·076 ± 0·001	1816·80 ± 0·10	10·20	2·253 ± 0·0048
11·20	3·065 ± 0·001	1717·12 ± 0·11	12·20	2·244 ± 0·0040	11·20	3·066 ± 0·001	1717·06 ± 0·20	12·18	2·243 ± 0·0048
13·20	1·067 ± 0·001	1625·03 ± 0·10			13·15	1·068 ± 0·001	1625·00 ± 0·08		
10·15	4·071 ± 0·001	1766·04 ± 0·05	11·12	2·240 ± 0·0048	10·15	4·070 ± 0·000	1765·92 ± 0·10	11·15	2·235 ± 0·0044
12·10	2·093 ± 0·001	1671·80 ± 0·12	13·00	2·208 ± 0·0048	12·15	2·092 ± 0·001	1671·90 ± 0·14	13·02	2·212 ± 0·0044
13·90	0·298 ± 0·001	1592·70 ± 0·20			13·90	0·298 ± 0·001	1592·73 ± 0·06		
14·05	4·000 ± 0·001	1589·10 ± 0·06	15·05	2·193 ± 0·0033	14·05	4·002 ± 0·001	1589·20 ± 0·08	15·05	2·191 ± 0·0035
16·05	1·934 ± 0·001	1504·76 ± 0·08	16·82	2·142 ± 0·0033	16·05	1·936 ± 0·001	1504·91 ± 0·09	16·82	2·144 ± 0·0035
17·60	0·448 ± 0·001	1448·76 ± 0·06			17·60	0·449 ± 0·001	1448·81 ± 0·07		
13·40	4·337 ± 0·001	1615·12 ± 0·07	14·40	2·209 ± 0·0038	13·40	4·340 ± 0·001	1615·10 ± 0·08	14·40	2·205 ± 0·0035
15·40	2·337 ± 0·001	1531·03 ± 0·08	16·30	2·157 ± 0·0038	15·40	2·338 ± 0·001	1531·10 ± 0·06	16·30	2·158 ± 0·0035
17·20	0·542 ± 0·001	1461·80 ± 0·08			17·20	0·543 ± 0·001	1461·85 ± 0·08		
17·15	4·392 ± 0·001	1465·43 ± 0·11	18·12	2·091 ± 0·0040	17·15	4·393 ± 0·001	1465·26 ± 0·10	18·12	2·083 ± 0·0043
19·10	2·405 ± 0·001	1395·29 ± 0·08	20·10	2·078 ± 0·0040	19·10	2·405 ± 0·001	1395·37 ± 0·10	20·10	2·080 ± 0·0043
21·10	0·465 ± 0·001	1331·37 ± 0·09			21·10	0·465 ± 0·001	1331·35 ± 0·08		
17·95	4·552 ± 0·000	1438·62 ± 0·05	18·95	2·091 ± 0·0026	17·95	4·551 ± 0·001	1438·51 ± 0·05	18·95	2·088 ± 0·0030
19·95	2·550 ± 0·001	1369·62 ± 0·05	20·95	2·069 ± 0·0026	19·95	2·550 ± 0·000	1369·65 ± 0·08	20·95	2·071 ± 0·0030
21·95	0·547 ± 0·000	1305·56 ± 0·08			21·95	0·547 ± 0·000	1305·51 ± 0·05		
21·00	4·208 ± 0·001	1332·26 ± 0·14	21·85	1·985 ± 0·0054	21·00	4·207 ± 0·001	1332·30 ± 0·06	21·88	1·986 ± 0·0045
22·70	2·458 ± 0·001	1280·06 ± 0·03	23·72	1·976 ± 0·0054	22·75	2·458 ± 0·001	1280·10 ± 0·11	23·75	1·975 ± 0·0045
24·75	0·463 ± 0·001	1224·04 ± 0·14			24·75	0·462 ± 0·001	1224·08 ± 0·05		
25·15	5·630 ± 0·001	1210·56 ± 0·12	26·10	1·991 ± 0·0056	25·15	5·614 ± 0·001	1209·72 ± 0·05	26·10	1·978 ± 0·0058
27·05	3·710 ± 0·001	1159·97 ± 0·11	28·05	1·929 ± 0·0056	27·05	3·699 ± 0·001	1159·64 ± 0·14	28·05	1·924 ± 0·0058
29·05	1·756 ± 0·000	1112·69 ± 0·05			29·05	1·743 ± 0·000	1112·46 ± 0·03		
28·00	4·879 ± 0·001	1139·13 ± 0·11	29·00	1·928 ± 0·0056	28·00	4·880 ± 0·001	1139·05 ± 0·08	29·00	1·929 ± 0·0040
30·00	2·866 ± 0·001	1091·61 ± 0·09	31·00	1·929 ± 0·0056	30·00	2·868 ± 0·001	1091·55 ± 0·05	31·00	1·923 ± 0·0040
32·00	0·854 ± 0·001	1046·76 ± 0·05			32·00	0·854 ± 0·000	1046·80 ± 0·06		
29·00	3·899 ± 0·001	1115·46 ± 0·04	30·00	1·926 ± 0·0022	29·00	3·898 ± 0·001	1115·46 ± 0·10	30·00	1·926 ± 0·0046
31·00	1·820 ± 0·000	1067·82 ± 0·02			31·00	1·820 ± 0·000	1067·85 ± 0·05		
32·00	3·480 ± 0·000	1046·06 ± 0·03	32·95	1·917 ± 0·0038	32·00	3·480 ± 0·000	1046·15 ± 0·05	32·95	1·916 ± 0·0043
33·90	1·460 ± 0·000	1003·73 ± 0·08			33·90	1·460 ± 0·000	1003·85 ± 0·08		
34·00	5·392 ± 0·000	1001·28 ± 0·02	34·70	1·860 ± 0·0040	34·00	5·392 ± 0·000	1001·42 ± 0·03	34·70	1·870 ± 0·0042
35·40	3·895 ± 0·000	972·26 ± 0·03	36·12	1·836 ± 0·0040	35·40	3·896 ± 0·000	972·23 ± 0·04	36·12	1·836 ± 0·0042
36·85	2·390 ± 0·000	944·55 ± 0·03	37·60	1·806 ± 0·0040	36·85	2·390 ± 0·000	944·51 ± 0·02	37·60	1·801 ± 0·0042
38·35	0·881 ± 0·000	918·27 ± 0·08			38·35	0·880 ± 0·001	918·27 ± 0·08		
38·10	5·131 ± 0·000	922·73 ± 0·07	38·85	1·803 ± 0·0040	38·10	5·123 ± 0·000	923·00 ± 0·03	38·85	1·803 ± 0·0041
39·60	3·608 ± 0·001	897·07 ± 0·03	40·32	1·796 ± 0·0040	39·60	3·600 ± 0·000	897·32 ± 0·02	40·30	1·792 ± 0·0041
41·05	2·194 ± 0·000	874·17 ± 0·03	41·80	1·777 ± 0·0040	41·00	2·196 ± 0·000	874·63 ± 0·03	41·78	1·773 ± 0·0041
42·55	0·673 ± 0·000	850·66 ± 0·03			42·55	0·673 ± 0·000	851·13 ± 0·04		

Table 1—*continued*

Heating sequence					Cooling sequence				
Temperature (° C)	*Beckmann reading (corrected)*	*Time of flow (s) corrected*	*Mean temperature* (° C)	*(ε′/k)/1000 (not corrected for density variation)*	*Temperature* (° C)	*Beckmann reading (corrected)*	*Time of flow (s) corrected*	*Mean temperature* (° C)	*(ε′/k)/1000 (not corrected for density variation)*
42·45	5·150 ± 0·000	852·09 ± 0·02			42·40	5·150 ± 0·000	852·14 ± 0·06		
			43·18	1·755 ± 0·0045				43·15	1·758 ± 0·0038
43·90	3·653 ± 0·001	830·00 ± 0·04			43·90	3·652 ± 0·000	830·00 ± 0·03		
			44·65	1·748 ± 0·0045				44·65	1·742 ± 0·0038
45·40	2·165 ± 0·001	808·87 ± 0·04			45·40	2·155 ± 0·001	808·81 ± 0·02		
			46·15	1·754 ± 0·0045				46·18	1·748 ± 0·0038
46·90	0·647 ± 0·001	788·00 ± 0·05			46·95	0·646 ± 0·001	788·12 ± 0·02		
46·95	5·136 ± 0·001	788·12 ± 0·03			46·92	5·148 ± 0·001	788·26 ± 0·04		
			47·68	1·754 ± 0·0038				47·65	1·764 ± 0·0040
48·40	3·678 ± 0·001	768·75 ± 0·02			48·40	3·680 ± 0·001	768·65 ± 0·04		
			49·12	1·702 ± 0·0038				49·12	1·709 ± 0·0040
49·85	2·196 ± 0·001	750·28 ± 0·03			49·85	2·196 ± 0·000	750·08 ± 0·102		
			50·60	1·680 ± 0·0038				50·60	1·680 ± 0·0040
51·35	0·716 ± 0·001	732·67 ± 0·04			51·35	0·717 ± 0·001	732·48 ± 0·03		
51·35	5·267 ± 0·001	733·71 ± 0·02			51·35	5·268 ± 0·001	733·53 ± 0·04		
			52·12	1·683 ± 0·0041				52·12	1·690 ± 0·0045
52·90	3·770 ± 0·000	716·43 ± 0·02			52·90	3·768 ± 0·001	716·15 ± 0·04		
			53·62	1·667 ± 0·0041				53·65	1·668 ± 0·0045
54·35	2·284 ± 0·001	699·98 ± 0·04			54·40	2·282 ± 0·001	699·69 ± 0·02		
			55·12	1·654 ± 0·0041				55·15	1·661 ± 0·0045
55·90	0·773 ± 0·001	683·92 ± 0·02			55·90	0·777 ± 0·001	683·63 ± 0·02		
56·20	5·192 ± 0·001	681·10 ± 0·00			56·20	5·190 ± 0·001	681·15 ± 0·04		
			56·95	1·656 ± 0·0052				56·95	1·660 ± 0·0046
57·70	3·691 ± 0·001	665·74 ± 0·02			57·70	3·686 ± 0·001	665·70 ± 0·02		
			58·45	1·635 ± 0·0052				58·45	1·652 ± 0·0046
59·20	2·197 ± 0·000	651·09 ± 0·06			59·20	2·184 ± 0·001	650·82 ± 0·03		
			59·95	1·642 ± 0·0052				59·95	1·645 ± 0·0046
60·70	0·703 ± 0·001	636·83 ± 0·03			60·70	0·702 ± 0·001	636·66 ± 0·04		
60·90	5·323 ± 0·001	635·75 ± 0·03			60·80	5·418 ± 0·001	636·90 ± 0·07		
			61·65	1·594 ± 0·0050				61·60	1·595 ± 0·0056
62·40	3·785 ± 0·001	621·99 ± 0·03			62·40	3·848 ± 0·001	622·84 ± 0·02		
			63·08	1·565 ± 0·0050				63·10	1·577 ± 0·0056
63·75	2·432 ± 0·001	610·45 ± 0·04			63·80	2·366 ± 0·001	610·10 ± 0·03		
			64·58	1·592 ± 0·0050				64·60	1·600 ± 0·0056
65·40	0·910 ± 0·001	597·61 ± 0·02			65·40	0·874 ± 0·001	597·46 ± 0·03		
65·40	5·242 ± 0·001	597·08 ± 0·03			65·45	5·240 ± 0·001	596·78 ± 0·04		
			66·12	1·586 ± 0·0068				66·18	1·611 ± 0·0043
66·85	3·852 ± 0·001	585·74 ± 0·05			66·90	3·842 ± 0·002	585·21 ± 0·02		
			67·58	1·593 ± 0·0068				67·60	1·588 ± 0·0043
68·30	2·406 ± 0·001	574·23 ± 0·02			68·30	2·408 ± 0·001	573·84 ± 0·02		
			69·30	1·596 ± 0·0068				69·30	1·570 ± 0·0043
70·30	0·474 ± 0·001	559·32 ± 0·05			70·30	0·478 ± 0·001	559·20 ± 0·03		
70·20	4·755 ± 0·001	560·08 ± 0·04			70·20	4·765 ± 0·001	560·41 ± 0·04		
			71·00	1·537 ± 0·0066				71·00	1·547 ± 0·0066
71·80	3·010 ± 0·001	547·53 ± 0·02			71·80	2·997 ± 0·001	547·60 ± 0·04		
			72·60	1·558 ± 0·0066				72·60	1·558 ± 0·0066
73·40	1·500 ± 0·000	536·85 ± 0·05			73·40	1·465 ± 0·001	536·75 ± 0·03		

Note: two values of the standard deviation of $(\epsilon/k)/1000$ have been calculated for each group of readings taken with any one setting of the Beckmann thermometer (covering a temperature range of 4 to 5° C), one for heating and the other for the cooling sequence.

Table 2. *Measured activation energy for thrice-distilled water in the range* 6–72° C

											Standard deviation for the group
Temperature (° C)	6·22	8·12	10·20	11·14	12·19	13·01	14·40	15·05	16·30	16·82	
ϵ/k/1000	2·3500 ± 0·0020	2·3395 ± 0·0055	2·2525 ± 0·0005	2·2375 ± 0·0025	2·2435 ± 0·0005	2·2100 ± 0·0020	2·2070 ± 0·0020	2·1920 ± 0·0010	2·1575 ± 0·0005	2·1430 ± 0·0010	0·0023
Temperature (° C)	18·12	18·95	20·10	20·95	21·86	23·74	26·10	28·05	29·00	30·00	
ϵ/k/1000	2·0870 ± 0·0040	2·0895 ± 0·0015	2·0790 ± 0·0010	2·0700 ± 0·0010	1·9855 ± 0·0005	1·9755 ± 0·0005	1·9845 ± 0·0065	1·9265 ± 0·0025	1·9285 ± 0·0005	1·926 ± 0·00000	0·0026
Temperature (° C)	31·00	32·95	34·70	36·12	37·60	38·85	40·31	41·79	43·16	44·65	
ϵ/k/1000	1·9260 ± 0·0030	1·9165 ± 0·0005	1·8650 ± 0·0050	1·8360 ± 0·0000	1·8035 ± 0·0025	1·8030 ± 0·0000	1·7940 ± 0·0020	1·7750 ± 0·0020	1·7565 ± 0·0015	1·7450 ± 0·0030	0·0024
Temperature (° C)	46·16	47·66	49·12	50·60	52·12	53·64	55·14	56·95	58·45	59·95	
ϵ/k/1000	1·7510 ± 0·0030	11·7590 ± 0·0050	1·7055 ± 0·0035	1·6800 ± 0·0000	1·6865 ± 0·0035	1·6675 ± 0·0005	1·6575 ± 0·0035	1·6580 ± 0·0020	1·6435 ± 0·0085	1·6435 ± 0·0015	0·0039
Temperature (° C)	61·62	63·09	64·59	66·15	67·59	69·30	71·00	72·60			
ϵ/k/1000	1·5945 ± 0·0005	1·5710 ± 0·0060	1·5960 ± 0·0040	1·5985 ± 0·0125	1·5905 ± 0·0025	1·5830 ± 0·0130	1·5420 ± 0·0050	1·5580 ± 0·0000			0·0071

Temperature (° C)	0	10	20	30	40	50	60	70
Correction $= T^2\beta/1000$	$-0·004_2$	$0·007_0$	$0·017_6$	$0·027_7$	$0·037_7$	$0·047_7$	$0·058_0$	$0·068_2$
Δ (*correction*) *for* 10° C	$0·011_2$	$0·010_6$	$0·010_1$	$0·010_0$	$0·010_0$	$0·010_3$	$0·010_2$	

Table 3. *Data on the steps observed with thrice-distilled water between 6 and 72° C*

Ordinal no. of flat	1	2(*a*)	2(*b*)	3	4	5	6	7	8	9	10
$\frac{\epsilon'}{k}$/1000	2·350	2·248	2·205	2·084	1·983	1·926	1·804	1·748	1·675	1·643	1·593
Temperature at middle of flat (° C)	6 ± 1	11·2	14·3	19·4	24·2	30·3	39·0	45·6	52·6	58·3	65·6

above-mentioned values of the energy change $\Delta\epsilon$ at each step, it is nevertheless worth investigating how far the energies of the molecular rotational and vibrational modes would be reflected in the activation free energy ϵ. It is expected that the more refined experiments now in hand on ethylene glycol, ethyl alcohol and other similar liquids will help to elucidate these aspects of the problem.

The plotted points for steps Nos. 2 and 5 in the inset to Fig. 5 show major deviations from the straight line, each amounting to 2 deg C, and because these deviations are nearly five times the estimated accuracy of the temperatures given in Table 3, it appears probable that in this region there are two series of steps superposed on each other. Thus the one series of steps would occur at equal intervals of nearly 7 deg C, and the other with a shorter repetition interval of about 5·5 deg C, so that the two sets overlap exactly at about 0, 20 and 40° C, and are partially resolved at intermediate temperatures. This would also serve to explain the different values of the temperature interval in Figs 6(*a*) and (*b*). It

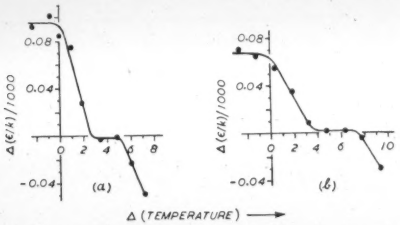


Fig. 6. Averaged shape of a step, (*a*) in the range 6 to 35° C, and (*b*) in the range 35 to 70° C.

would therefore be expected that, with a smaller measuring interval, $\Delta T = 1$ deg C, one of the steps in the neighbourhood of 30° C would also be resolved into two steps separated by 2 to 3 deg C. These detailed experiments are in hand and will be reported separately.

6. Comparison with previous standard viscosity data

At this stage, it is possible to compare the foregoing results with the values of ϵ/k calculated from the standard viscosity data on water given in BS 188: 1937 (based on the data collected by Bingham and Jackson 1918–19), which data are believed to be accurate to 1 part in 5000 on a relative scale. These calculated values of $(\epsilon/k)/1000$ are plotted in the upper half of Fig. 5 (open circles), and the graph exhibits several well-defined segments, the limits of which are indicated by the lower set of vertical arrows, but does not show any of the sharp drops to be found in the experimental graphs drawn in the lower part of Fig. 5. This lack of detail is largely attributable to the fact that Bingham and Jackson's values of viscosity are all smoothed over a range of the order of 5 to 10 deg C, which would naturally obscure all detail in temperature intervals smaller than this figure. Therefore a proper comparison with Bingham and Jackson's data can only be made if the measuring interval ΔT, which is 2 deg C in the present experiments, is increased by a factor of 3 or 4.

Such an increase to 6 deg C is readily effected by taking the average of three readings differing successively by 2 deg C, and the corresponding smoothed graph is plotted as solid circles at the top of Fig. 5, with a displacement of 0·1 unit above the graph based on Bingham and Jackson's data. The agreement between these two curves is quite satisfactory, and may be taken as proof that Bingham and Jackson's values of viscosity are consistent with the more accurate measurements given in Tables 1 and 2. Also, the increased detail already obtained with the measuring interval of $\Delta T = 2$ deg C serves to emphasize the need for refined measurements with an even smaller interval of 1 deg C. It is hoped to publish a table of accurate values for the viscosity of water (relative to that at 20 deg C) when these measurements are completed. Similar refined measurements are also being made on ethylene glycol and glycerol in order to obtain an explanation for the more gradual drop between the energy steps in these two liquids.

Acknowledgments

The authors are indebted to Dr. Salimuzzaman Siddiqui for his constant interest and encouragement and to several colleagues for help during preparation of the manuscript.

References

ANDRADE, E. N. DA C., 1934, *Phil. Mag.*, **17**, 497, 698.
BINGHAM, E. C., and JACKSON, R. F., 1918–19, *Bull. Nat. Bur. Stand.*, **14**, 59.
BROCKHOUSE, B. N., 1957, *Acta Cryst.*, **10**, 827.
CROSS, P. C., BURNHAM, J., and LEIGHTON, P. A., 1937, *J. Amer. Chem. Soc.*, **59**, 1134.
GLASSTONE, S., LAIDLER, K. J., and EYRING, H., 1941, *The theory of rate processes*, 480–516 (New York: McGraw-Hill Book Co.).
NASEEM, M., and QURASHI, M. M., 1958, *Pakistan J. Sci. Industr. Res.*, **1**, 197.
QURASHI, M. M., 1958, *Pakistan J. Sci. Industr. Res.*, **1**, 116.
RAUF, A., and QURASHI, M. M., 1959, *Pakistan J. Sci. Industr. Res.*, **2**, 30.
TAYLOR, H. S., and GLASSTONE, S., 1951, *A treatise on physical chemistry*, **2**, 3rd ed. (London: Macmillan and Co. Ltd.).

Shaft position indicator for use with journal bearings

by J. J. HUNTER, Fluids Division, and C. J. HUGHES, Lubrication and Wear Division, National Engineering Laboratory, Glasgow

MS. received 22nd June 1960

Abstract

A description is given of equipment using transistors and designed for the indication and measurement of the position of the shaft within the bush of a journal bearing system. Shaft movements cause variations in the impedance of inductive transducers, unbalancing the a.c. bridge of which they are part. The bridge output, in the form of an amplitude-modulated carrier, is supplied to an a.c. amplifier and then demodulated to recover a voltage proportional to shaft displacement. By monitoring the bearing on two diameters at right angles and supplying the demodulator outputs to the X- and Y-plates of a cathode-ray oscilloscope, a polar plot of shaft movement is possible and the form of vibration can be observed directly.

Although designed for use on a bearing of about 1 *in. diameter, the equipment is easily adjustable for other bearing sizes and can give useful information with shaft speeds up to* 20 000 *rev/min.*

Circuit details are given and discussed and an assessment is made of the accuracies of the various units.

Introduction

In most problems of hydrodynamic lubrication it is, in theory, possible to predict the behaviour of any particular system. The prediction involves a knowledge of some six variables and their interaction, and solutions usually make use of a number of assumptions and approximations. In bearings, the parameter of vital importance is the distance separating the moving and stationary surfaces; to help to correlate theoretical and empirical design it is necessary to measure this distance.

In the case of a plain journal bearing, i.e. a smooth cylindrical shaft rotating in a smooth hollow cylinder, this involves the measurement of the position of the shaft under various conditions of speed, load and lubricant supply. In certain circumstances the shaft can vibrate in the bush, so the problem is not merely one of measuring a static deflection, but includes dynamic deflections and, possibly, their phase relationship with a point (e.g. an unbalanced load) on the rotating shaft. The magnitude of the deflection varies with the size of the bearing, but it is usual practice to make the bore of the bush 1/1000 to 3/1000 in. per in. larger than the shaft diameter, e.g. a nominal 1 in. diameter bearing would have a 'clearance' between 0·001 and 0·003 in.

Various workers have tackled the problem in the past. Direct contact methods (Clayton and Jakeman 1936) have been used with very low shaft speeds, but results are questionable where hydrodynamic films can be formed under the sensing elements. Optical systems (Newkirk and Grobel 1934) require high-resolution lenses and present a number of difficulties in phase indication. Air-gauging (Barwell, F. T., and Hughes, C. J. 1953, *Lubrication Division Report L.D.R.* 22/53, National Engineering Laboratory, unpublished) is simple but has a poor frequency response.

Use can be made of the electrical impedance variations produced in a suitable transducer by journal/bush displacements, and many methods apply this in one form or another. It has even been found worth while to use the resistance changes (Siripongse, Rogers and Cameron 1958, El-Sisi and Shawki 1958) which occur in the hydrodynamic oil film, but for consistent results high currents, which may damage surfaces, are necessary and calibration is difficult. Generally, systems in which shaft/bush displacements change capacitive (Dayton and Simons 1951) or inductive (Greenhough 1948) couplings have been the most popular. Capacitive systems present problems of dielectric variation when they include either the lubricating film or air outside the bearing, in the first instance because of pressure effects and cavitation of the oil film, and in the second because of contamination of the air gap with escaping lubricant. Inductance methods are adversely affected by varying magnetic conditions in the shaft or external magnetic fields.

However, because of experience of the stability of inductance systems it was decided, after investigation of the magnetic homogeneity of typical shaft materials, to adopt this type of transducer and to continue with the development, for this purpose, of a high-speed measuring system which had already proved successful in indicating displacements in oil hydraulic machinery (Hunter, J. J., and Leslie, W. H. P., *Transistorized General Purpose Equipment for Mechanical Measurements from zero frequency to* 1 *kc/s*, National Engineering Laboratory Fluids Report, in preparation).

Requirements

Part of the N.E.L. programme of research on hydrodynamic bearings involves the investigation of bearing behaviour under conditions of dynamic load. For this work, a small machine has been made with test bearings about 1 in. in diameter and shaft speeds variable between 4000 and 20000 rev/min. Using typical clearance ratios, diametral clearances may be up to 0·003 in., and thus any measuring system must be capable of measuring displacements appreciably less than this. Both shaft and bush can be measured separately to an accuracy of the order of 0·00001 in. and it is desirable that dynamic displacement measurements should have a similar accuracy, although a greater error might be acceptable. At a speed of 20000 rev/min the equipment should be capable of resolving movements ocurring within a fraction of a millisecond, if an adequate amount of information is to be obtained within one revolution.

To these essential requirements may be added others designed to facilitate the use of equipment and the rapid interpretation of results. Attachment to the test bearing

should not be a complicated process and calibration should be simple. Results should be presented in a manner easy to comprehend; ideally, the measuring equipment should amplify actual shaft movements so that the display reproduces shaft behaviour. Unfortunately thermal distortions in the bearing may confuse simple measurements and conflicting requirements occur. For example, to demonstrate the general behaviour of the bearing, a simple picture of shaft eccentricity, ignoring distortions as much as possible, is desirable, but for more detailed studies of the hydrodynamics, actual film thickness could be useful. The apparatus should therefore be capable of both functions.

Design

To obtain a complete picture of the position of the shaft within the bush, it is necessary to measure in a minimum of two planes spaced along the axis of the bearing; this assumes that there is no geometric distortion of the bearing components. Each plane requires measuring stations on at least two co-ordinates, and thus four measuring stations are required for the description of the shaft position at any instant. This implies four independent channels of information which can be correlated for display purposes. The channels can be similar. If the measuring planes of the bearing are at right angles to the axis of the system and the co-ordinates are at right angles to each other, the measuring stations are disposed in a similar fashion to the plates of a cathode-ray tube. This leads to a convenient form of displacement display that has, in fact, been utilized by previous workers. The output signals from each measuring station are supplied to the appropriate X- and Y-plates of a cathode-ray tube, causing the spot to deflect, its motion reproducing the motion of the axis of the shaft in the particular measuring plane.

The arrangement adopted in the present work is shown in the block diagram (Fig. 1), the inductive transducers being

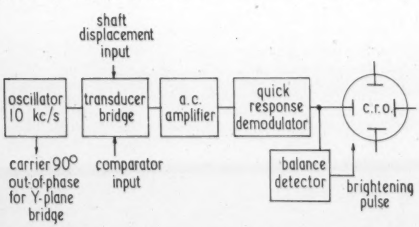


Fig. 1. Block layout of system used for the shaft position indicator.

arranged in an a.c. bridge. The output from the bridge, which is a carrier-frequency signal, amplitude-modulated by the transducers, is then amplified and finally demodulated to recover a signal directly proportional to displacement. To enable the shaft movement to modulate the carrier, the shaft is used as part of the magnetic circuit of the inductive transducer, as shown in Fig. 2. For this type of circuit the relationship between transducer inductance and air gap is of the form

$$L = A + \{B/(C + d)\}$$

where L is the inductance, A the stray shunt inductance, B a function of the windings in the transducer, C a constant proportional to the iron path and d is the air gap. As part of the iron path is in the shaft, which may be rotating, tests have been made which show that the material used is magnetically homogeneous and causes no significant variation in the constant C. It can be seen from this relationship that the total air gap must be large compared with the expected change in air gap if a nearly linear displacement/output relationship is desired. In general, the relationship was sufficiently linear if the change in air gap was 10% or less of the total air gap (Fig. 3). An improvement in linearity is to be expected if the transducers are used in a 'push–pull' arrangement, since the change in inductance is the difference of the changes in two similar transducers and non-linearities will tend to cancel. The change in inductance with air gap determines the system gain which must be used, following the transducer, to provide a reasonable output signal for display and recording. The total inductance, and therefore transducer winding, is decided by the loading that can be put on the bridge voltage supply.

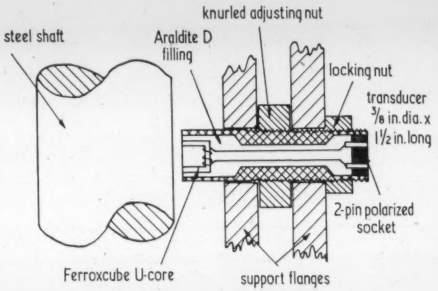


Fig. 2. Inductive displacement transducer.
Core: made by splitting Ferroxcube E core (Mullard FX 1052/A4).
Winding: 900 turns of 46 s.w.g. enamelled copper wire.
Inductance: approximately 20 mH, d.c. response approximately 95 Ω.

Details of windings, inductance and core are given in Fig. 2. Mechanically, the transducers have been designed to facilitate interchange, each transducer being mounted in a brass case and provided with a polarized socket connection (Fig. 4). A fine screw thread on the outside allows for easy mounting and adjustment of the initial air gap. For general display of a locus of the shaft centre, a pair of transducers is arranged in a half-bridge circuit for each diameter. An advantage of this arrangement is that if the transducers are adjusted to be equal in inductance with the shaft centred, a symmetrical change in air gap or change in the bridge supply voltage does not change the zero or balance point of the bridge, since equal changes in the transducers cancel out. The only error at points other than zero is due to the change in sensitivity with change in total gap due, say, to thermal expansion. For an initial total gap of 0·015 in. and a range of 0·0015 in., a change in initial total gap to 0·0135 in. gives an output 18% higher for the same range of 0·0015 in. (Fig. 5), instead of the 100% error that would have resulted if a single transducer had been used. The unbalance of the bridge, produced by shaft displacement, can be measured by using a reference, or comparator, transducer pair or the secondary of the bridge supply transformer as a ratio arm.

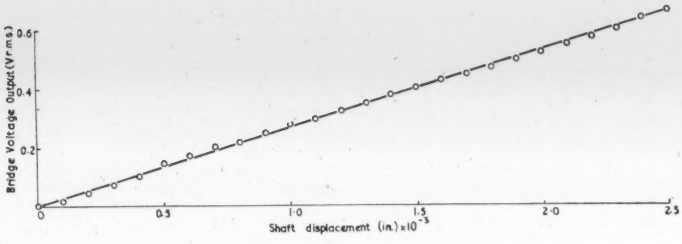


Fig. 3. Transducer characteristic.
Initial gap: 0·015 in. Bridge supply: 6 v r.m.s.

In the circuit shown in Fig. 6, with the switch S_1 in its centre position, the transformer centre tap is earthed. The transducers can now be arranged to be precisely equal in inductance and thus, if they are similar, to have equal air gaps. In position 2, the switch 'earths' the bridge on to a potential divider across two 10% tappings on either side of the centre tap of the transformer secondary. This allows

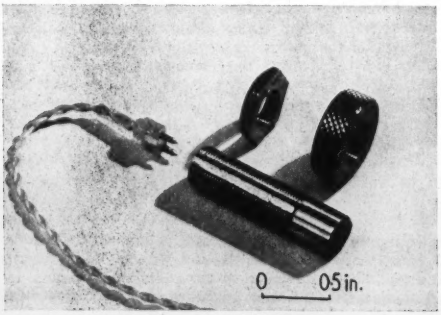


Fig. 4. Inductive transducer.

the ratio arm to be adjusted to compensate for any expected variations of the measuring transducer. There are substantial advantages in stability to be gained by using mutually coupled inductive ratio arms in this fashion (Clark and Vanderlin 1949), as the bridge is relatively insensitive to stray earth impedances. In the third switch position the bridge is 'earthed' at the centre point of a similar push–pull transducer which is on a 'comparator' stand. In this 'comparator' (Fig. 7), a dummy shaft is used and X- and Y-plane transducers are disposed as in the test rig. The shaft, which is in the form of a lever, is deflected by a pair of micrometers acting along the same X- and Y-diameters as the transducers; the arrangement allows the small displacements necessary at the transducers to be made with more precision, as the micrometer movement is made greater by the lever ratio. With similar transducers in the test rig

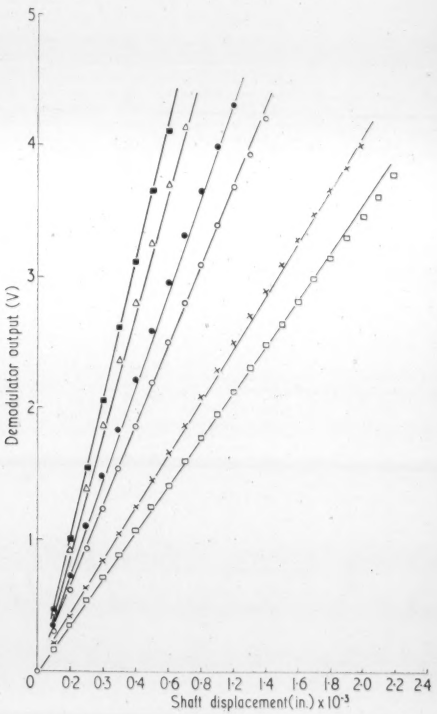


Fig. 5. Demodulator output voltage *versus* shaft displacement.
Transducer initial gap settings:
□ = 0·025 in. ● = 0·0135 in.
× = 0·0225 in. △ = 0·010 in.
○ = 0·015 in. ■ = 0·005 in.
a.c. amplifier gain = 100.
Bridge supply voltage = 6 v r.m.s.

and comparator, any micrometer adjustment at the comparator necessary to balance movements of the test shaft can be taken as a measurement of that movement multiplied

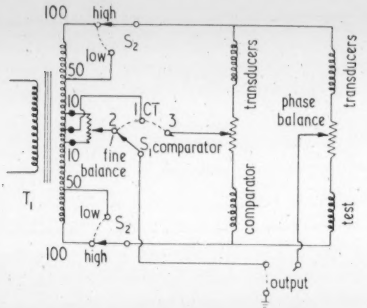


Fig. 6. Transducer bridge circuit.

T_1 = core, Ferrox cube 2 × FX 1105/A4 primary, 200 turns, 38 G copper secondary, 200 turns, 38 G copper. Tapped at centre and ten and fifty turns either side of centre.

by the lever ratio. If the transducers are not sufficiently similar, a calibration curve can be made, or the lever ratio can be altered. The magnitude of both static and recurrent dynamic displacements can be estimated in this matter.

If measurements of bush distortion, involving a change in total clearance on a diameter, are required, a simple re-arrangement allows a single test transducer to be balanced against a single transducer on the comparator, using the supply secondary as the ratio arm. For example, if one transducer at the end of a diameter on the test shaft is balanced by one X-transducer on the comparator, and the transducer at the opposite end of the test shaft diameter is balanced by one Y-transducer on the comparator, the sum of the deflections indicated by the X- and Y-comparator micrometers is an indication of total bearing clearance. If the displacements are varying, each can be displayed on the separate beams of a twin-beam cathode-ray oscilloscope and summed at any instant.

As indicated in Fig. 1, the bridge of each channel is supplied with a stable voltage at 10 kc/s and the bridge output, modulated by shaft movement, is amplified in an a.c. amplifier. This amplified signal is then demodulated, giving a voltage, directly proportional to shaft movement, which is supplied to an oscilloscope. For each plane an oscilloscope with two similar d.c. amplifiers, for driving the X- and Y-plates, is required. These d.c. amplifiers need not be of very high gain or stability, as the demodulated output signal can be up to ± 5 v and provision is made for a stable reference mark, indicating bridge balance, to be superimposed on the trace. To produce the reference mark, the demodulator output signal is also connected to a circuit, which produces an output when the demodulator output voltage lies within a band of 20 mv about the zero reference level. This 'marker' output, indicating bridge balance, can be applied to an oscilloscope beam-brightening circuit, so that the oscilloscope spot, tracing out a path (Fig. 8) directly related

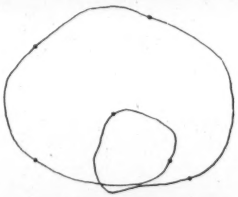


Fig. 8. Typical locus trace.

to the movement of the shaft centre, is brightened whenever it crosses the predetermined X- and Y-axes. The trace can be moved about in relation to these axes, and measured by adjusting the comparator micrometers. For permanent records the trace can be photographed or, if detailed analysis is required, the output from each channel can be recorded separately using a multi-channel recorder.

If correlation between the locus trace and shaft rotation is desired, the output from a suitable transducer, working against a reference mark on the shaft, can be used to mark the trace, the mark taking the form of a bright spot (of different intensity to the 'marker' spot) or a line at right angles to the trace. Additionally, the bridge balance mark can be used to drive a stroboscope and, if there is a simple relationship between locus and shaft rotations, a reference mark on the shaft can be viewed directly.

The system can be calibrated statically, if all transducers have similar characteristics, by balancing the test transducers against those on the comparator, the balance being indicated conveniently by the beam brightness. If the transducers are not similar, a comparative calibration is necessary, and may be made by moving the test shaft in known increments and noting the deflections of the comparator shaft needed to produce bridge balance.

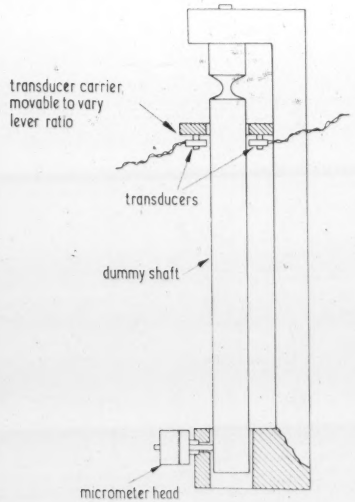


Fig. 7. Diagram of comparator. The system is similar in a plane at right angles to the paper along the axis of the shaft.

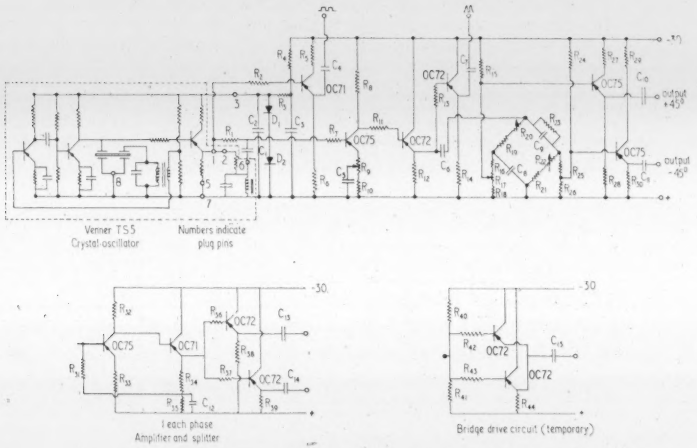


Fig. 9. Circuit diagram of oscillator supply system.

Circuit details and characteristics

Oscillator and bridge supply (*Fig.* 9). A commercial 10 kc/s crystal oscillator of high-frequency stability is used to supply a stable sine wave output voltage derived from a square wave voltage, *via* a filter tuned to the fundamental. The amplitude of the square wave depends directly on the oscillator supply voltage, which is stabilized by voltage reference diodes (D_1 and D_2). The sine wave is supplied to a phase-splitting network and the outputs from this network are then amplified, giving two bridge supply voltages which are 90° out-of-phase with respect to each other. To avoid mutual interference, the X- and Y-transducer bridges, in one plane, are supplied from separate quadrature voltages as the phase-sensitive demodulator ignores signals produced by voltages 90° out-of-phase.

A.C. amplifier (*Fig.* 10). This is basically a three-stage

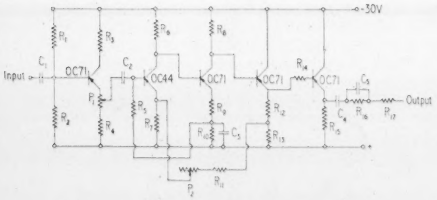


Fig. 10. Circuit diagram of a.c. amplifier.

feedback amplifier with a gain of 100 and buffer stages at input and output. The overall gain can be varied continuously by a potential divider following the input buffer stage. Owing to the use of a large amount of overall feedback, the amplifier has a highly stable gain and low distortion. The gain/frequency characteristic (Fig. 11) has been shaped to be suitable over the frequency band centred around 10 kc/s. The gain has fallen sharply by 50 c/s, thus minimizing the possibility of mains interference.

In addition to the use of feedback to minimize noise, the first stage transistor of the amplifier is run at low operating current and voltage, thus keeping the input transistor noise low. Measurements of equivalent noise input voltages showed that the noise level of the basic amplifier is equivalent to an input signal not greater than 15 μv. The distortion at maximum output, measured in terms of the second harmonic,

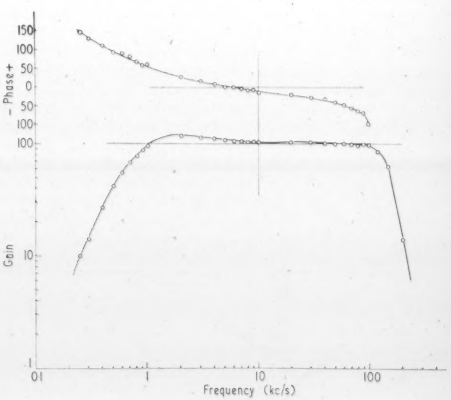


Fig. 11. Amplifier characteristic.

is only 0·17% of the fundamental, the third harmonic content being immeasurable. At a 1 v r.m.s. output the distortion is less than the noise. The equivalent noise input for the whole amplifier unit is 120 μv. Distortion of a 1 kc/s signal at 20 v peak-to-peak output produces a second harmonic of

0·4% of the fundamental, and a third harmonic of 0·25%. Higher harmonics were below the noise level. At the normal maximum working output of 10 v peak-to-peak, the second harmonic is 0·03% and the third harmonic 0·018% of the fundamental.

Demodulator (*Fig.* 12). This is a transistor, phase-sensitive demodulator (Hunter, J. J., *A transistor demodulator*, to be published) which samples the a.c. input signal once per cycle and then holds the sample voltage for the rest of the cycle, until the next sampling point. Originally, the demodulator was designed for full-wave operation, as the boxed transistors in Fig. 12 indicate. However, difficulties in matching transistors led to an unacceptable amount of ripple and the use of the full-wave circuit has been deferred until more suitable transistors are available. The demodulator has a low d.c. drift; tests have shown drifts of less than 1 mv/h and less than 5 mv over a period of 24 h, at an ambient temperature around 19° C. The output impedance of the stage is about 2·5 kΩ, making it suitable for supplying following transistor stages. The maximum open circuit output voltage range is ±5 v.

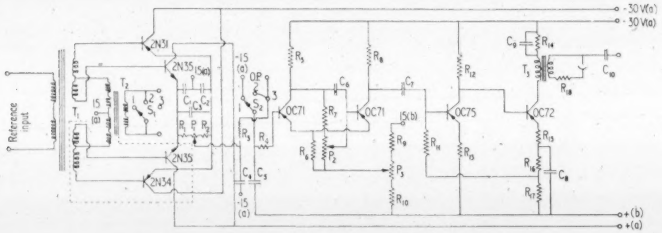


Fig. 12. Circuit diagram of demodulator and marker.

T_1 = six windings of 100 turns, filar wound 38 s.w.g. Inductance 32 mH. Core: 2 Ferroxcube FX1105/A4.

T_2 = Wound as T_1. Connected: Primary, two windings. Each half secondary, two windings.

T_3 = Primary: 400 turns, 46 s.w.g. 60 mH. Secondary: 400 turns, 46 s.w.g. 60 H. Core: 2 Ferroxcube FX1052/A4.

Balance marker (*Fig.* 12). The balance-marker circuit (Hunter, J. J., *A balance-marker circuit*, to be published) is basically a high-gain amplifier, with a narrow working band about 20 mv wide, and with the feedback arranged to make it oscillate whilst the input lies within this band. The working band can be centred on the reference voltage, and temperature tests have shown the zero point to have a drift of only 10 mv over a temperature range of 20° C. This working band and stability represent a resolving accuracy of 0·3% of the demodulator full-scale output of ±5 v. The oscillation frequency of the balance marker is about 50 kc/s and its amplitude is about 100 mv peak-to-peak; it is amplified before the output. The amplifier also drives a neon indicating lamp, which permits initial alignment of the system or static measurements without an oscilloscope. A neon lamp has been used as an indicator since the original proposal was to use it as a simple stroboscope. A modification is being considered as the neon lamp has shown hysteresis effects.

Power supply. A 36 v bank of accumulators provides a supply free from mains interference and from fluctuations over a period of hours. (Tests show a drift of less than 0·1 v over 1 h and a drop of less than 0·2 v over 8 h.) The initial output voltages are controlled by a power transistor, the base voltage of which can be adjusted by a variable potential divider.

Application and use of the equipment

In the design of the equipment emphasis was laid on easy application and calibration, but there are two main points where care must be exercised. They are:

(*a*) the adjustment of the transducers so that the bridge is balanced for some known position of the shaft within the bush;

(*b*) the setting of the demodulator and marker circuits to zero position.

The first is necessary because the initial bridge balance position is the reference point from which all readings are taken. To make full use of the circuit design the transducers should be adjusted with the test shaft concentric with the bush, but, because of the practical difficulty of fixing the test shaft in this concentric position, it is easier to set the transducers by using a shaft which fits the bush exactly. The position of the transducers is fixed by setting one with a feeler gauge, and then adjusting the others until balance is obtained with the feeler removed. The final adjustment of bridge balance, for in-phase and quadrature components, can best be carried out by monitoring the bridge on a cathode-ray oscilloscope, and for this purpose test points are provided on the bridge panel.

The second point, the setting of demodulator and marker references, is accomplished by the operation of switch S_1, S_2, and the variable resistors P_1 and P_3 in Fig. 12, which are accessible from the front panel. In the first position of the switch, the marker circuit input is shorted to the reference voltage. The marker is adjusted by P_3 into the centre of its working band, as indicated by the beam brighteners of the cathode-ray tube or by the lighting of the indicator lamp. This circuit, now set at zero, can be used to align the demodulator. The switch, turned to position 2, connects the marker to the demodulator, but shorts the demodulator input, allowing the demodulator output to be set to zero by potentiometer P_1. In position 3, the normal operating position, the demodulator input is connected to the a.c. amplifier and the marker circuit indicates when the signal component in phase with the reference is zero. Once the initial balance of the quadrature component, as indicated on the cathode-ray oscilloscope, has been obtained, readjust-

ment is usually unnecessary and the marker circuit can be used to indicate bridge balance. The disadvantage of having any appreciable amount of quadrature component is that the total signal the amplifier can handle is limited. Any increase in the quadrature component decreases the maximum output available from the demodulator, as only the in-phase component of the signal provides an output.

Performance

Accuracy. Although no exhaustive long-term tests have been made to ascertain the limits of accuracy, the main causes of possible error have been examined and operating experience to date has shown no significant variations.

Accuracy must be evaluated taking into account the measuring technique employed.

Recurrent phenomena. When measuring recurrent phenomena the equipment can be used as a comparator, that is, to compare the test displacement with a displacement produced by a reference micrometer. The main factors which can cause errors are:

(*a*) the cross-coupling of the X- and Y-planes at the transducers;
(*b*) the stability of the test and reference transducers;
(*c*) the stability of the demodulator zero; and
(*d*) the stability of the zero and the resolution of the balance marker.

The cross-coupling factor (*a*) has two components. One is the change in gap caused by a curved surface moving across the transducer pole-faces, i.e. movement of the shaft in a direction perpendicular to that being measured. The other is magnetic cross-coupling. Tests have shown that the combined effect of these produces error outputs of the order 50 mv maximum for shaft movements within the working range, which represents a possible error of 0·5% of full-scale output. Item (*b*) can also be considered in two parts; the mechanical stability of the transducers and their electrical stability. The mechanical stability, in practice, includes the stability of the shaft and bush and is primarily affected by temperature changes. A typical temperature change of 30° c may cause a transducer movement of 0·0001 in. giving rise to an error of 2% at most. A systematic error of this type can easily be allowed for, as the total gap can be measured at the operating temperature. Electrical stability can be affected by change, due to heating, in the permeability of the Ferroxcube core or shaft, and change in resistance of the transducer winding. Permeability is only a second-order factor of the transducer inductance characteristic, because the reluctance of the air gaps exceeds that of the core and shaft and, to a first approximation, change in resistance only affects the magnitude of the quadrature component which is ignored in the demodulator. In any case, these can only produce errors if the two transducers of a pair are affected in a dissimilar manner, and it is unlikely that large temperature gradients will exist across the system. Drift in the demodulator zero (*c*) is small and only likely to cause errors of the order of 0·05%. The 'marker' circuit has a resolution and stability of 0·3% of the full-scale output of ±5 v. The errors in items (*c*) and (*d*) would, of course, be proportionately greater for smaller output voltages.

Non-recurrent phenomena. For non-recurrent phenomena the scale factor or magnification of the equipment is also important. This depends mainly on the stabilities of:

(*a*) the bridge supply voltage;
(*b*) the transducer sensitivity (as dealt with in item (*b*) under recurrent-phenomena);
(*c*) the amplifier gain; and
(*d*) the demodulator gain.

Tests have shown that the voltage supplied to the bridge (*a*) only varies 50 mv over a period of 12 h, producing a possible error of 0·5% of full scale. A combination of bridge supply voltage variation and a change in amplifier gain gave drifts in the amplifier output not greater than ±25 mv, which was the accuracy of the recorder used for the tests. This stability was maintained over two days during which period the temperature varied about 4° c. It represents a variation band of 0·5% of full scale.

Speed of response. Theoretically, the 10 kc/s carrier and the quick response demodulator are capable of providing ten discrete points of information in a millisecond. This is adequate for a system in which no significant change can occur in less than a tenth of a millisecond, and can indicate even higher rates of change if the information can be built up over a number of repetitive cycles. The amplifier has a gain/frequency response suitable for a bandwidth of a few kilocycles. As an exact correlation between time and frequency response is not simple, it is the intention to carry out an overall calibration of the equipment. A step or ramp function can be fed to the transducers by rotating a specially machined shaft at various speeds, and the phase response of the system can be measured by using the balance marker output to drive a stroboscope and illuminate a scale attached to the shaft.

Conclusions

The equipment is at present being used successfully to measure shaft behaviour in a bearing subjected to a combination of steady and rotating load. Its application has proved simple and easily accomplished by personnel without specialist electronic knowledge, even to the extent of their being able to recognize and detect the source of spurious signals. The unit chassis construction has aided development work and will aid maintenance, although up to the present this has been negligible. It will, of course, be realized that the technique is applicable to the measurement of any small displacement, and individual channels can be used independently if a separate oscillator carrier source is available.

During the bearing investigation work a number of refinements have been suggested. At present, the information from successive cycles of the shaft rotation under 'whirling' conditions tends to be superimposed during the display. Using a phonic wheel to indicate shaft angular positions, and normal batch counting techniques for selection, it would be possible to display any required portion of the vibration trace. For continuous portrayal of successive revolutions, a form of three-dimensional 'helical' display could be arranged by adding extra voltages from a time base to the inputs of both the oscilloscope amplifiers. The measurement of shaft position at both ends of a bearing enables the direction of shaft axis to be calculated at any instant. However, if the output of the transducer pairs in any axial plane were compared continuously, it would be possible to display the axial tilt. In such a display, conical whirl would be easily recognizable, whereas the individual monitoring at the bearing ends requires a phase-relationship indicator to provide this information.

Acknowledgments

The work described has been carried out in the National Engineering Laboratory of the Department of Scientific and Industrial Research. This paper is published by permission of the Director.

The authors wish to acknowledge the help of Mr. W. H. P. Leslie, the equipment having been constructed in the instrumentation section of Fluids Division under his direction, and of Mr. J. Bruce, who carried out noise tests on the amplifier. Mr. J. A. Cole was associated with the initiation of this work.

References

CLARK, H. A. M., and VANDERLIN, P. B., 1949, *Proc. Instn Elect. Engrs* (II), **96**, 365, 386.

CLAYTON, D., and JAKEMAN, C., 1936, *Proc. Instn Mech. Engrs*, **134**, 437.

DAYTON, R. W., and SIMONS, E. M., 1951, *N.A.C.A. TN*2544.

EL-SISI, S. I., and SHAWKI, G. S. A., 1958, *Amer. Soc. Mech. Engrs*, Paper No. 58-A-253.

GREENOUGH, M. L., 1948, *Trans Amer. Inst. Elect. Engrs*, **67**, 589.

NEWKIRK, B. L., and GROBEL, L. P., 1934, *Trans Amer. Soc. Mech. Engrs*, **56**, 607.

SIRIPONGSE, C., ROGERS, P. R., and CAMERON, A., 1958, *Engineering* (*London*), **186**, 146.

CORRESPONDENCE

Dynamic measurement of elasticity using resonance methods

Mr. G. Bradfield (1960) draws attention to the use of a variety of magnetostrictive transducers for measuring the elastic properties of various materials, including coal. These methods, while presumably suitable for a detailed study of single specimens, would not appear to be practicable where large numbers must be examined; for example, the wrapping of a thin wire round the edges of small specimens, as in Figs C and H, would seem to be a very tedious operation.

The flaws and inhomogeneities that exist in coal and other natural materials make it essential to use large numbers of specimens in order to obtain representative values of the elastic moduli, and the scatter of individual values renders superfluous the highest possible accuracy of measurement. What is required is a method that is capable of rapid application and for which the apparatus is simple and inexpensive. Such a method is that of the composite oscillator where the transducer is a nickel rod (Terry and Woods 1954, 1955). This was very successfully used by the late Dr. N. B. Terry at the Mining Research Establishment for measuring the elastic properties of coal. The cementing process is not difficult, and to an experimental accuracy of about ½%, which is sufficient for this purpose, the precise adjustment of specimen length to that of the transducer is not required and the error due to the cement is negligible (Terry 1957).

National Coal Board,
Mining Research Establishment,
Worton Hall,
Worton Road,
Isleworth,
Middx.

I. EVANS
25*th October* 1960

The letter of Mr. I. Evans refers to a method of measuring elastic constants which we have used, among numerous others, many years ago but which we abandoned for the simpler methods we describe. We have tested hundreds of specimens using the latter methods and the fitting of the specimen, contrary to Mr. Evans' suggestion, only takes a fraction of a minute; cementing a specimen takes considerably longer and, in addition, the use of a hot cement is undesirable and generally necessitates a waiting period for cooling.

Basic Physics Division,
National Physical Laboratory,
Teddington,
Middx.

G. BRADFIELD
30*th November* 1960

BRADFIELD, G., 1960, *Brit. J. Appl. Phys.*, **11**, 478.
TERRY, N. B., 1957, *Brit. J. Appl. Phys.*, **8**, 270.
TERRY, N. B., and WOODS, H. J., 1954, *Proc. Leeds Phil. Soc.*, **6**, 251.
——, 1955, *Brit. J. Appl. Phys.*, **6**, 322.

The influence of the method of demagnetization on the permeability in soft magnetic materials

Ten years ago I began to make some measurements on the low field permeability of soft magnetic materials such as silicon iron and Permalloys. I observed that their permeability depends on the manner by which they are demagnetized (Smolinski 1952).

Heating above the Curie temperature gives lower values of low field permeability than those obtained after a gradual reduction of an alternating field. This effect appeared in four materials on which the experiments were performed:

(i) Silicon steel—4% silicon, normal metallurgical process (0·02% carbon).
(ii) Silicon steel—2·5% silicon, vacuum melted (0·01% carbon).
(iii) Permalloy B—50% nickel, vacuum melted.
(iv) Permalloy C—76% nickel, vacuum melted.

The specimens were in the form of cut bars 76 mm long, 10 mm wide and 0·35–0·4 mm thick, inserted in a Permalloy yoke provided with a coil of 400 turns (Lamson 1940, Smolinski 1950). The permeability was measured on an Owen type a.c. bridge (Sullivan type AC 1100/1102) with a wave analyser (Radiometer FRA 1) as null-point indicator.

The following procedure was adopted. The specimens were first heated above the Curie temperature. They were then inserted in the coil and the permeability measured, (1) with an alternating field increasing from zero up to one to two oersteds; (2) with the alternating field gradually decreasing to zero, and (3) as in (1). The samples were then taken away from the yoke and demagnetized by pulling them slowly out of a coil carrying a large alternating current. The specimens were re-inserted in the yoke and permeability was measured (4) as the alternating field increased from zero, and (5) as in (4).

The results of measurements are shown in Figs 1–4 for the investigated materials in the form of a relation

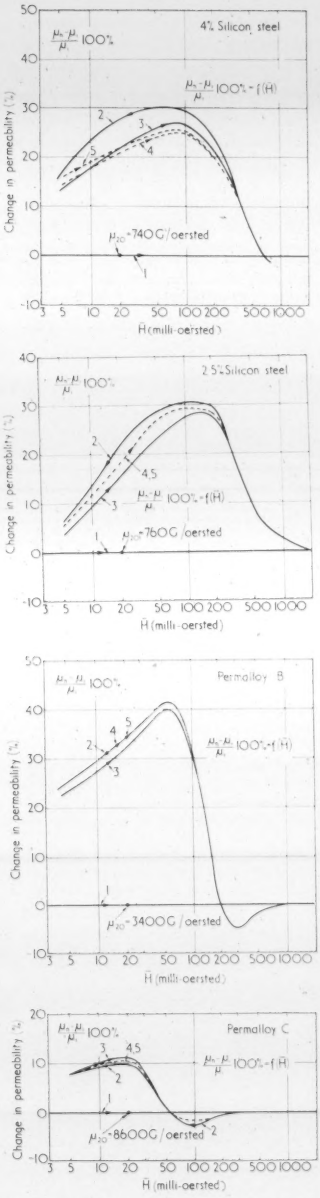
$$(\mu_n - \mu_1)/\mu_1\ 100\% = f(\bar{H})$$

where the subscripts 1–5 and the figures on the graphs correspond to the five processes described above. These curves exhibit a maximum of from 10 to 40%, occurring between 80 and 150 milli-oersteds.

These results show clearly that the low and medium field permeability is always greater after a.c. demagnetization than after thermal demagnetization. This was confirmed later, for instance in the papers of Sommerkorn (1955), and of Jackson, Lee and Troughton (1958).

It should be pointed out that there is sometimes a small minimum in the curves (e.g. in Figs 1, 3 and 4) which means that higher field permeability may be greater after thermal demagnetization than after a.c. demagnetization.

When repeating these experiments, it was found that it is not necessary to heat the specimens above the Curie point to obtain their thermal demagnetization. Temperatures below 200° C were sufficient to exercise this effect (Smolinski 1953). In these experiments the specimen was in the form of stamped rings of 4% silicon iron, 0·35 mm thick, 68 mm external diameter, 40 mm internal diameter, stacked to a height of 15 mm. After annealing it was cooled slowly through the Curie point to room temperature. Each ring was then insulated from the other and the whole specimen was provided with a winding. The permeability at 20 milli-



Figs 1–4. Percentage change in permeability after various demagnetizing processes (described in text).

oersteds was measured as before (point μ_c, Fig. 5), and the sample was demagnetized by pulling it through the demagnetizing coil carrying a large alternating current as before. To

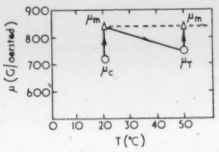


Fig. 5. The method of measuring the permeability μ_T. For explanation of symbols see text.

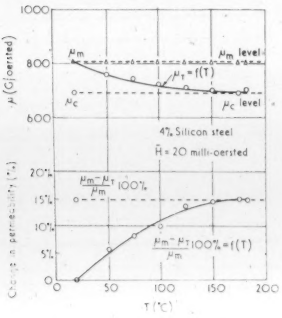


Fig. 6. 4% silicon steel, $\overline{H}$ = 20 milli-oersteds.

avoid the effect of disaccommodation the following measurements were done on the next day. First the low field permeability μ_m was measured and found to be greater than μ_c. Then the specimen was heated to 50° c and cooled to room temperature, 20° c. The low field permeability was then found to be μ_T, such that $\mu_c < \mu_T < \mu_m$ (Fig. 5).

The whole process was repeated, the temperature T being raised steadily until μ_T reached a constant value equal to μ_c. This appeared for the 4% silicon–iron at about 170° c (Fig. 6). It should be pointed out that after a.c. demagnetization (and relaxation), the permeability μ_m was constant.

Nearly the same results were obtained with Permalloy B (Smolinski 1954). These results show clearly that thermal demagnetization of these materials appears at a temperature much lower than the Curie temperature. It is considered that a satisfactory explanation of this effect is not yet known.

Technical University of Warsaw,
Warszawa,
Poland.

A. SMOLINSKI
22nd November 1959

REFERENCES

JACKSON, R. C., LEE, E. W., and TROUGHTON, A. G. H., 1958, *Brit. J. Appl. Phys.*, **9**, 495.
LAMSON, H., 1940, *Proc. Inst. Radio Engrs*, 541.
SMOLINSKI, A., 1950, *Prace Badawcze Panstwowego Instytutu Telekomunikacyjnego* (2), 77.
—— 1952, *Arch. Elektrotech. (Warsaw)* (1), 67.
—— 1953, *Arch. Elektrotech. (Warsaw)* (3/4), 337.
—— 1954, *Instytut Podstawowych Problemów Techniki PAN, Pracownia Magnetyczna*, Internal Report (6).
SOMMERKORN, G., 1955, *Tech. Mitt. Krupp* (4), 71.

Notes and comments

New journals

During the past few months we have received copies of several new journals which may be of interest to readers. They are as follows:

Cryogenics

This journal deals specifically with low temperature engineering and research. It appears quarterly, publishing original papers on all aspects of cryogenic research and technology, together with shorter technical notes, book reviews and correspondence. It also contains a bibliography of recent cryogenic literature, together with 4-language summaries.

Cryogenics is edited jointly by Dr. K. Mendelssohn, of the Clarendon Laboratory, Oxford; R. B. Scott, chief of the N.B.S. Cryogenic Engineering Laboratory, Colorado; and Professor L. Weil of the Institut Fourier, Grenoble University. In addition, a panel of twelve experts act as advisory editors.

The first issue contains a contribution by Dr. N. Kurti, who discusses the prospects of nuclear cooling. There are also nine original papers from Britain, America, France and the Netherlands.

In addition to the work devoted to fundamental issues, *Cryogenics* will also publish papers dealing with the increased use of cryogenic propellants as rocket fuels, new methods of handling, storage and transport, low temperature electronics and other technical problems related to cryogenics.

The journal is published by Heywood and Co. Ltd., Drury House, Russell Street, Drury Lane, London, W.C.2. The annual subscription is £5.

Electronics weekly

A new weekly publication, in newspaper format, was launched on 6th September 1960 by the publishers, Messrs. Heywood and Co. Ltd. It is intended to provide topical news of interest to all people concerned with the electronics industry and includes news of a scientific, technical, commercial, economic, financial and statistical nature. The publication will endeavour to maintain an international flavour in order to keep readers informed of trends in the industry throughout the world.

Electronics weekly is published every Wednesday price 4*d*. Further information can be obtained from the publishers at Drury House, Russell Street, Drury Lane, London, W.C.2.

Journal of Nuclear Energy: Part C

We received the first issue of the *Journal of Nuclear Energy: Part C* which combines Numbers 1 and 2 of Volume 1. This new journal, edited by Dr. J. V. Dunworth, covers plasma physics, accelerators and thermonuclear research. It is intended to publish translations of the numerous and important papers produced in Russia in order to make them available for physicists unable to read the originals. Among the contributions contained in the first issue are the following: M. A. Leontovich: Some investigations on the physics of a heated plasma and its containment; J. W. Burren, D. Morgan and W. Walkinshaw: Production of high density proton currents for colliding beam accelerators; Mitsuo Imoto: Essential limitations and *TN-t* diagram for self-sustaining fusion systems; J. D. Lawson: On the classification of electron streams; P. Reynolds and H. M. Skarsgard: A plasma betatron; B. Lehnert: On the confinement of charged particles in a magnetic field; E. W. Laing, S. J. Roberts and R. T. P. Whipple: Equilibrium configurations of a toroidal plasma; P. F. Smith: An experimental and theoretical investigation of the magnetic field produced by ridged pole-pieces.

The journal is published quarterly and the subscription is £7 per volume (U.S.A. $20). The publisher is Pergamon Press Ltd.

Nuclear engineering abstracts

We have received a new abstracting journal published under the above title which, as its title suggests, deals mainly with nuclear engineering subjects. Its purpose is to present and record developments and trends reflected in the current technical literature and it is mainly designed for engineers, scientists, manufacturers and industrialists either engaged in or connected with the nuclear engineering field.

The abstracts are collected together under one or more of the following main subject headings: *Science and engineering*—General; *Fission:* Research techniques, instrumentation and testing methods, Plant and auxiliaries, Subsidiary processing and handling; *Fusion:* Plant and auxiliaries, Subsidiary processing and handling; *Radiation*, Special associated subjects.

The Editor is Mr. H. L. J. Burgess, former Editor of *Vacuum*, and the journal is published by Silver End Documentary Publications, Ltd., 9–11, Tottenham Street, W.1. The journal is issued quarterly and the annual subscription is £6 6*s*. ($18.50), single copies £1 12*s*. 6*d*. ($4.75).

Optics and Spectroscopy

We have received from the Optical Society of America a copy of the Society's translation of the Russian journal, *Optika i Spektroscopiva*. This monthly journal of the USSR Academy of Sciences commenced publication in 1956, and publishes the work of leading Russian scientists in all branches of optics and spectroscopy, including x-ray, ultraviolet, visible infra-red and microwave, thin layer optics, filters, detectors, diffraction gratings, electroluminescence, thermal radiation backgrounds, infra-red polarizers and their many other applications in science and industry. The English translation, which is being published with a grant-in-aid from the National Science Foundation, starts from Volume 6, January 1959, and it is the aim of the Society that *Optics and Spectroscopy* will appear within four months of the Russian original.

The subscription rate is $28, including postage, and further information can be obtained from the Society, 1155 16th Street, N.W., Washington 6, D.C., U.S.A.

The Physics of Fluids

The American Institute of Physics has announced that its journal, *The Physics of Fluids*, will in future appear monthly, instead of bimonthly, effective with Volume IV (1961).

By doubling the number of issues it is hoped that it will be possible to arrange for prompter reporting of the results of research on the structure, statistical mechanics and general physics of gases, liquids, plasmas and other fluids. As before, the journal will continue to offer papers on certain

basic aspects of physics of fluids bordering geophysics, astrophysics and related fields.

Further information can be obtained from the American Institute of Physics, 335 East 45th Street, New York 17, N.Y., U.S.A.

Radioactive sources and chemicals

Two catalogues have recently been issued by the Radiochemical Centre of the United Kingdom Atomic Energy Authority. *Radioactive chemicals*—deals with the primary isotopes and labelled compounds which are of interest to all who use radioactive tracer methods and with radioactive pharmaceuticals. It also includes radioactive reference sources for calibrating instruments. *Radioactive sources*—deals with sealed radioactive sources and appliances which will be of interest to radiotherapists and to industrial users of ionizing radiations. It also includes a great variety of radioactive sources which can be used for research purposes.

Both catalogues give the tariff for irradiation in the research reactors at Harwell. Most of the chemicals and sources in regular production are listed and users are recommended to choose these whenever possible on grounds of economy and ready availability: but the Centre welcomes inquiries for other materials or appliances which can often be made to special order and is willing to advise users as to the most suitable product for this particular purpose.

Only the briefest technical information is given in the catalogues but detailed specifications for most products can be supplied on request. The Centre's scientific staff will gladly advise on technical matters generally.

Copies of the catalogue are obtainable on request from the Radiochemical Centre, Amersham, Buckinghamshire.

Quantum electronics conference

The second international conference on quantum electronics will be held in Berkeley, California, March 23rd–25th, 1961. Emphasis will be placed on basic theory, progress, and new research efforts in the field. Topics will include methods for the generation of millimeter and shorter waves, coherent sources and amplifiers (high frequency masers, irasers, lasers), and fundamental studies of materials and techniques suitable for the higher frequencies. The conference is sponsored by the Office of Naval Research.

Attendance will be limited to those active in research; for further details write to Professor J. R. Singer, Chairman, Second International Conference on Quantum Electronics, Department of Electrical Engineering, University of California, Berkeley 4, California.

Editorial and publications sales offices

Readers are requested to note that all correspondence of an editorial nature should now be sent to the following address:

The Editor and Deputy Secretary,
The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7 (Kensington 0048).

Correspondence concerning subscriptions, missing journals, and sales of publications should be sent to:

The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1 (Belgravia 6111).

Journal of Scientific Instruments

Contents of the February issue

SPECIAL ARTICLES

Some developments and applications of the optical lever. By R. V. Jones.
Design of scales for industrial instruments (abstract). By A. J. Maddock.

PAPERS

Mass spectrometer for the isotropic analysis of lithium. By R. G. Ridley and D. E. P. Silver.
Transistorized vacuum trip unit. By D. Jones.
Simple constant-temperature laboratory crystallizer. By B. M. Bartlett.
Recording integrator for gas chromatography. By A. P. H. Jennings.

LABORATORY AND WORKSHOP NOTES

Improved manufacturing technique for blood-flow probes. By S. J. Sterrett.
Simple mercury expansion regulator for control of air temperature to ± 0·02 deg C. By K. L. Butcher and M. J. Byrne.
Modification to the model 540 scanning count integrator. By E. Mariniansky.
Welding of conducting electrodes to thin metallic films. By St. Dembicka-Jellonkowa.
Simple circulating pump. By A. Macfadyen.
Device for preparing accurately x-ray oriented crystals. By W. L. Bond.

CORRESPONDENCE

Daylight factor meter. From D. P. Turner.

NOTES AND NEWS

New instruments, materials and tools New books Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.
EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.
REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.
ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611–2.)
SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12*s.* 6*d.* each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.
CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

UNWIN BROTHERS LIMITED, WOKING AND LONDON

Galvanomagnetic effects and their application

by C. HILSUM, B.Sc., Ph.D., F.Inst.P., Services Electronics Research Laboratories, Baldock, Herts.

MS. received 5th *January* 1961

Abstract

The simple theory of Hall effect and magnetoresistance is discussed. The use of various semiconductors for galvanomagnetic applications is considered, and some practical applications described.

1. Introduction

Investigations of the basic physics of solids often include studies of the galvanomagnetic effects. The Hall effect, the voltage developed across a conductor when it is placed in crossed electric and magnetic fields, reveals the sign and concentration of the charge carriers, while the increase in resistance of the conductor, the transverse magnetoresistance effect, yields information on the interaction between the carriers and the crystal lattice. Though many physicists are aware of such academic applications, it is not generally realized that there now exist a number of practical devices which exploit these two effects.

In this survey we first discuss the simple theory of the Hall effect and magnetoresistance. We then consider the materials which can be used for these applications and show that certain compound semiconductors are particularly suitable. Finally we describe some of the instruments which have been developed.

2. The Hall effect

Electrons travelling in a crystal under the action of an electric field drift from one current electrode to the other. When a transverse magnetic field H is applied as in Fig. 1

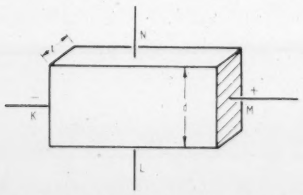


Fig. 1. Arrangement for observing the Hall effect.

the electrons are deflected, and a charge builds up on the lower side of the crystal. In a very short time the repulsive force caused by this charge balances the deflecting force caused by the magnetic field and the electrons once again drift directly along the crystal. In the new equilibrium a voltage V_H has been created across the crystal, and, since the repulsive force must be equal to the magnetic Lorentz force, we have

$$\frac{V_H e}{d} = veH \qquad (1)$$

where v is the drift velocity of the electrons, and d the width of the crystal. The drift velocity is assumed to be the same for all the carriers and at low electric fields it is proportional to the carrier mobility μ. The current density J is equal to nev, n being the electron concentration, so that

$$V_H = \frac{HJd}{ne} = \frac{HI}{t}\frac{1}{ne} \qquad (2)$$

where t is the thickness of the sample and I the current. The Hall voltage is proportional to the product of current and magnetic field, and inversely proportional to the sample thickness. The constant of proportionality is called the Hall coefficient R_H, and in this case it is equal to $-1/ne$.

A similar argument can be applied when the conduction takes place by means of positively charged carriers, such as holes in a p-type semiconductor. If the field directions are as shown in Fig. 1 the holes drift from right to left, so that in this case as well the carriers will be deflected by the magnetic field to the lower side of the crystal. The charge that builds up here is now positive, but again an equilibrium is reached when the repulsive electric force is equal to the Lorentz force, and the Hall voltage is given by Eqn (2), with the negative sign removed and the hole concentration p substituted for the electron concentration n.

The situation is quite different when the crystal contains appreciable numbers of both holes and electrons. In the simplest case, with the hole concentration equal to the electron concentration and the hole mobility μ_p equal to the electron mobility μ_n, the Hall voltage will be zero. The magnetic field diverts both types of carrier to the lower side of the crystal, and there they recombine. No repulsive field is created, and in equilibrium the carriers drift diagonally across the crystal. We can show that this has an important effect on the electrical resistance.

3. Magnetoresistance

Before the magnetic field is applied to the n-type semiconductor the electrons drift directly along the sample. A short time after the field has been applied, when the new equilibrium has been set up, the electrons again drift along the sample (Fig. 2). The resistance of the sample is therefore unchanged, since on our simple theory the drift velocity is independent of the magnetic field. Similarly the application of the magnetic field does not alter the resistance when the conduction takes place entirely by holes.

For the 'mixed' conductor with $n = p$ and $\mu_n = \mu_p$ the initial deviation of the carriers remains uncompensated, because no Hall electric field is created. In the equilibrium condition all the carriers travel at an angle across the sample. Since the drift velocity is unchanged, the forward component of drift velocity must be less than it was before the magnetic field was applied, and the resistance of the sample increases.

There will be few occasions when a conductor contains

equal concentrations of holes and electrons of equal mobility, but the general case can be solved without much difficulty. The sign and magnitude of the Hall field will be determined mainly by the more numerous or more mobile carriers. This field will tend to repel these carriers, and force them to travel more directly along the sample, but it must attract the

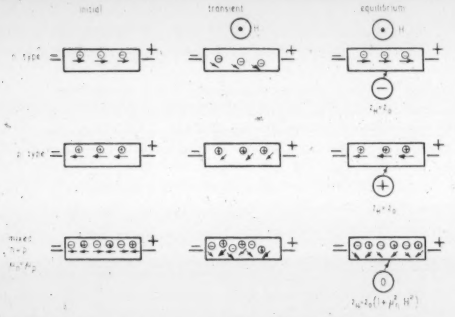


Fig. 2. Symbolic representation of carrier flow before and after application of the magnetic field (Z_0 = sample resistance in zero magnetic field, Z_H = sample resistance in magnetic field H).

carriers of the other sign and increase their deviation. The *total* current across the sample, in the direction of the Hall electric field, must be zero, but both the hole current and the electron current will have components across the sample, and the forward components of drift velocity will be less than they were before application of the magnetic field. The consequent increase in resistance is given in Table 1.

Table 1. *Expressions for Hall coefficient and magnetoresistance effect*

Type of semiconductor	*Hall coefficient, R_H*	*Magnetoresistance, Z/Z_0*
n	$-1/ne$	1
p	$+1/pe$	1
Mixed $\begin{cases} n = p \\ \mu_n = \mu_p \end{cases}$	0	$1 + \mu_n^2 H^2$
General case $n = pc$, $\mu_n = \mu_p b$.	$\dfrac{1 - b^2 c}{(1 + bc)^2} \dfrac{1}{pe}$	$1 + \mu_n^2 H^2 \dfrac{(b + c)/b(1 + bc)}{(1 - c)^2/(1 + bc)^2}$

The basic fact we have now established is that we get large magnetoresistance effects by reducing the Hall electric field. When the Hall field is small the carriers are free to travel at an angle to their original path, and their effective mobility is reduced. We might therefore deduce that if we take the n-type sample of Fig. 1, which shows no magnetoresistance, and connect the two Hall voltage electrodes L and N together (Fig. 3(*a*)), we should obtain an increase in resistance because we have suppressed the Hall electric field. Such an increase can indeed be observed. There are, however, more effective ways of reducing the Hall field and obtaining a magnetoresistance effect. One way is to use a sample shaped so that the width d of the current electrodes is much greater than the distance l between them (Fig. 3(*b*)). Such short wide samples have an inconveniently low resistance, but a structure equivalent to a number of them connected in series can be made (Welker 1955) by evaporating or plating conducting strips at intervals across a long thin sample of material (Fig. 3(*c*)). We can see that in this arrangement, called a raster-plate, transverse electric fields are short circuited. The structure in which the maximum resistance change is obtained is the Corbino disk, which has one electrode at the centre and the other at the circumference (Fig. 3(*d*)). No Hall electric field can be set up in such a sample, and the carriers spiral from one electrode to the other.

The derivation of the expressions given in Table 1 was based on the assumption that all the carriers of one sign had the same drift velocity. This is true for some semiconductors, but in others, notably germanium and silicon, there is a spread in the carrier velocities. It can be seen from Eqn (1) that under such conditions the Hall field could not exactly compensate for the Lorentz force for all the carriers. When the equilibrium is reached carriers with some average velocity will be travelling undeviated along the semiconductor. Faster carriers will be deflected towards the lower side of the crystal, because the Lorentz force on them will be stronger than the repulsive electric force. Slow carriers will experience too small a Lorentz force to compensate for the Hall field, and they will be deflected towards the upper side of the crystal.

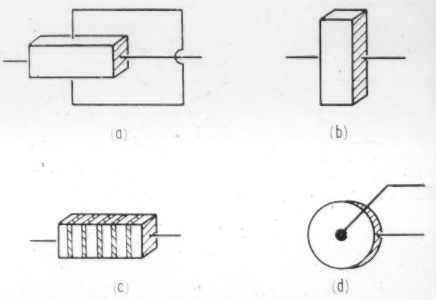


Fig. 3. Methods of increasing the magnetoresistance effect, (*a*) connected Hall electrodes, (*b*) a short wide sample, (*c*) a raster plate, (*d*) a Corbino disk.

These two groups of carriers have their forward component of drift velocity reduced by the application of the magnetic field, and the resistance of the crystal will increase. The magnitude of the magnetoresistance effect will depend on the spread in carrier velocities, and for small magnetic fields the resistance will increase by a factor $1 + F\mu^2 H^2$, where μ is the average mobility and F is determined by the velocity spread. Generally F is smaller than unity.

The magnetoresistance effects in any particular semiconductor sample will probably be due to a combination of the three processes we have outlined—the suppression of the Hall field because of the presence of both holes and electrons, the suppression of the Hall field by geometrical effects, and the decrease in effective mobility due to the spread in drift velocity.

4. Choice of material

It can be stated as a general rule that if a material is to be useful for galvanomagnetic applications the carrier mobility must be high and the carrier concentration low. The Hall coefficient is inversely proportional to the carrier concentration, and in a metal, for example, the concentration is so high that all the voltages developed are very small. This is one

reason why semiconductors are preferred. The necessity for a high carrier mobility in applications of magnetoresistance can be seen immediately from Table 1. The relative change in resistivity, $\Delta\rho/\rho_0$, for a mixed semiconductor involves terms in μ^2H^2, and similar relations apply when the carriers have a spread in drift velocity. The geometrical magnetoresistance effects are large when the deviation of the carriers in the magnetic field is large, and for this too a high mobility is needed.

For Hall effect applications it is important to consider the circuit which is connected across the Hall electrodes. When the Hall voltage is passed to an amplifier of high input impedance, the factor of merit is the Hall voltage divided by the power input to the crystal, and this is proportional to $R_H/\sqrt{\rho}$ (Ross, Saker and Thompson 1957). For this to be large the mobility must be high and the carrier concentration low. On the other hand, if the load across the Hall electrodes has an impedance chosen to match that of the crystal the important factor of merit is the transfer efficiency, or the power diverted to the load divided by the power input to the crystal. The transfer efficiency is proportional to $(R_H/\rho)^2$, which is equal to μ^2. The carrier concentration is in this case not directly relevant. A high efficiency is not, however, the only criterion of a useful device. Semiconductors are notoriously difficult to use because their electrical properties vary markedly with temperature. The mobility of a semiconductor is highest when all impurities have been removed, but, unfortunately, in this pure condition the carrier concentration is strongly temperature dependent. Doping, the intentional addition of impurities, will reduce the mobility, are normally made n-type, and the galvanomagnetic properties are then determined by the electron mobility. This simple procedure is not satisfactory for InSb. In principle InSb could be made n-type with a carrier concentration of 10^{18} cm^{-3} and an electron mobility near 20 000 cm^2/v sec, and the carrier concentration would be relatively temperature independent. Such material would have an inconveniently low resistivity. It is better to dope the indium antimonide slightly p-type, and the resistivity is then a little higher than in pure material. The Hall coefficient does not decrease steadily with temperature increase, but passes through a maximum at a temperature determined by the degree of doping. A useful operating temperature range can therefore be obtained by doping to put this maximum near room temperature (Hilsum 1960). Even this reduced temperature dependence is too large for many applications, and when sensitivity is not the prime consideration one can choose n-type InAs. For a still smaller temperature dependence and a more convenient resistivity n-type GaAs and n-type Ge might be recommended, but the mobility would be four or five times less than in InAs. A compromise can be made by using an 'alloy' of InAs and InP. The properties depend on the composition, and if only about 20% of the phosphide is included a mobility of nearly 10 000 cm^2/v sec is obtained, with a temperature dependence of carrier concentration of 0·04% per deg C.

We have noted that though a heavy doping of a low resistivity semiconductor like InSb can reduce the temperature dependent effects, the resistivity of the material becomes so low that no practical use can be made of it. Semiconductors

Table 2. *Electrical properties of some semiconductors suitable for galvanomagnetic applications*

Semiconductor	*Germanium*		*Silicon*		*Gallium arsenide*		*Indium arsenide*		*Indium antimonide*		*Mercury selenide*
	pure	*doped (n)*	*pure*	*doped (n)*	*pure*	*doped (n)*	*pure*	*doped (n)*	*pure*	*doped (p)*	*Evaporated film (n)*
Electron mobility (cm^2/v sec)	4000	3500	1500	1200	10 000	5000	33 000	23 000	78 000	37 000	3000
Hole mobility (cm^2/v sec)	2000	—	500	—	500	—	500(?)	—	750	600	—
Electron concentration (cm^{-3})	2.10^{13}	2.10^{15}	$\sim10^{10}$	10^{15}	$\sim10^{6}$	5.10^{16}	$\sim10^{15}$	5.10^{16}	$1{\cdot}6.10^{16}$	7.10^{16} (holes)	10^{17}
Resistivity (ohm cm)	50	1	2.10^{5}	5	6.10^{8}	0·025	0·2	0·005	0·005	0·033	0·02
Temp. dep. of Hall coeff. (%/deg)	−5	−0·01	−8	−0·1	−9	−0·01	−3	−0·06	−2	Total change for 20 deg C ±1·5%	−0·1
$R_H/\sqrt{\rho}$	4.10^{4}	3.10^{3}	10^{6}	3.10^{3}	2.10^{8}	8.10^{2}	10^{4}	2.10^{3}	6.10^{3}	5.10^{3}	4.10^{2}

but it also stabilizes the carrier concentration. In Table 2 are given the characteristics of a number of semiconductors which might be considered for galvanomagnetic applications.

It can be seen that the element semiconductors do not have as high an electron mobility as the compounds given. The hole mobility for the compounds is quite low, so that if the materials have to be doped to give temperature stability they are brittle, and samples cannot be rolled or worked to give very thin units. If devices could be made with a thickness of only a few microns, it would be practicable to employ heavily doped semiconductors. Thin films of HgSe have in fact been made with a mobility only a few times less than in large single crystal samples (Elpatevskaia and Regel 1956). The electrical characteristics of these films are included in

Table 2. Some success has also been achieved on InSb and InAs, and it is likely that in the near future it will be possible to use materials with a resistivity less than 10^{-3} ohm cm for galvanomagnetic applications.

5. Hall effect versus magnetoresistance

In some of the applications of the galvanomagnetic effects, it is possible to use either the Hall effect or magnetoresistance. We can be guided in our choice between them by a consideration of the magnetic field available. The Hall voltage is proportional to the magnetic field, whereas the magnetoresistance effect is proportional to the square of the field. At some particular field the two effects will be equally useful, and at lower fields than this the Hall effect should be chosen. The particular magnetic field at which this crossover occurs depends on the material, the shape of the sample, and the electrical load to which the sample is connected. For a square sample of InSb working into a matched load, the crossover field is 2000 oersteds, and if the load is large compared with the sample resistance, the crossover field is 500 oersteds. The crossover field is roughly inversely proportional to the mobility, so that usually the only materials for which one finds practical use for magnetoresistance effects are InAs and InSb.

6. Applications of the Hall effect

Some applications which make use of the Hall effect are listed in Table 3.

Table 3. *Applications of the Hall effect*

Application	*Current supply to Hall unit*	*Source of magnetic field*
Gauss meter	Local battery	Magnet
Compass	Local battery	Earth
Magnetometer	Oscillator	Stray fields
Susceptibility meter	Oscillator	Magnet
Clip-on ammeter	Local battery	Current in a wire
Measurement of ion and electron currents	Oscillator	Beam of charged particles
Wattmeter		
Modulator	Unknown current (a.c., d.c.)	Current in coil (a.c., d.c.)
Multiplier		
Amplifier		
Oscillator		
Microwave power measurement	Electromagnetic wave, electric vector	Electromagnetic wave, magnetic vector

In the simplest device, an instrument for measuring magnetic fields, a thin rectangular plate of the semiconductor has two current electrodes and two Hall effect electrodes attached. The plate, often called a 'Hall unit' or 'Hall generator', is supplied with current from a small dry battery, and a low resistance milliammeter is connected across the Hall terminals. The reading of the milliammeter will be dependent on the magnetic field in which the unit is placed, and InSb or InAs units can measure fields down to 1 oersted.

The sensitivity can be increased if the unit is placed between two permalloy rods, which concentrate the magnetic flux (Fig. 4). The gain depends on the ratio of the length of the rods to their diameter. With two 12 in. rods of $\frac{1}{4}$ in. diameter, an open circuit Hall voltage approaching a millivolt is obtained from an InSb unit in a magnetic field of 0·02 oersted (Ross, Saker and Thompson 1957). The system acts as an electrical compass, for the orientation of the rods in the earth's magnetic field is indicated on a sensitive meter which measures the Hall voltage.

Further improvement in sensitivity can be gained by amplifying the Hall voltage. The driving current is now

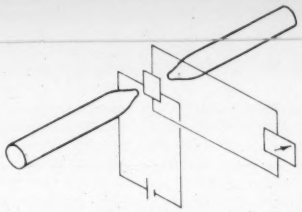


Fig. 4. The Hall effect compass.

obtained from a transistor oscillator, and the Hall voltage amplified by a transistor amplifier tuned to the oscillator frequency. Using InSb, fields down to 10^{-3} oersted can easily be detected without concentrators, and with rods 4 in. long and $\frac{1}{4}$ in. diameter the detection limit approaches 10^{-6} oersted (Hilsum and Thompson 1959, unpublished). The performance of this magnetometer is still not as high as that which is theoretically attainable. Since InSb shows no current noise (Oliver 1957), the limit to detection should be set by Johnson noise within the Hall generator, and fields down to 10^{-6} oersted should be detectable without concentrators, if the amplifier has a bandwidth of 1 c/s. In practice the limit is set by thermal drifts. InAs units can be used for magnetometers and, with 4 in. $\times$ $\frac{1}{4}$ in. concentrators, can detect fields less than 10^{-5} oersted (Weiss 1960). But in this application InSb is preferable. It is true that in InAs the Hall coefficient is relatively independent of temperature, so that the calibration of the magnetometer would be constant, but it is more important that zero drift should be small. The magnetometer is originally set so that a zero reading is shown on the output meter. No Hall generator is made with the two Hall probes exactly opposite each other, and the misalignment voltage must be compensated in some way. This is done either by connecting a variable resistance between one Hall electrode and one current electrode, or by supplying a small magnetic field from a permanent magnet or a solenoid. The zero adjustment is made when the instrument is brought into use, and it is obviously desirable that the balance should not change during the experiment. The zero drift, which is due to slight changes in temperature, is found to be less with InSb than with InAs (Hilsum and Thompson 1959, unpublished). A theoretical study (Voeikov 1958) indicates that the stability should be proportional to the carrier mobility, so that InSb should be about three times better than InAs on this count.

The Hall generator can be considered as a sensitive detector of magnetic flux. It is therefore analogous to a galvanometer and it can be used in a magnetic circuit in just the same way as a galvanometer can be connected into an electrical circuit. For example, the resistance balancing circuit of Fig. 5(*a*) is similar to a Wheatstone bridge with two equal batteries. Its equivalent is the magnetic circuit of Fig. 5(*b*), where the two air gaps replace the two resistances, soft iron bars replace the wires, and the magnet with a centre tap takes the place of the batteries and resistances in the arms ABC. In the balanced condition no flux passes along B′D′, but if the

reluctance of either air gap is changed, the balance is upset and a Hall voltage will be developed in the unit. The field in each gap can be near 2000 oersteds, and since the unit with the oscillator–amplifier combination can detect a field of 10^{-3} oersted, a change of reluctance of one part in a

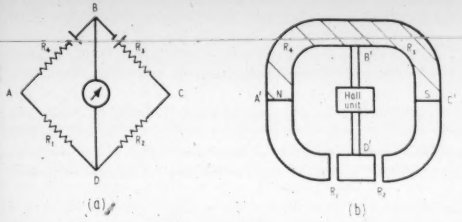


Fig. 5. The susceptibility meter, the magnetic analogue of the Wheatstone bridge.

million is observable (Hilsum and Rose-Innes 1958). The permeability of the material in the gap is $1 + 4\pi\chi$, χ being the volume susceptibility, so that the susceptibility of a material with $\chi = 10^{-7}$ c.g.s. unit can be measured. This magnetic circuit is therefore a system for measuring magnetic susceptibility and, though it does not have the sensitivity of the Gouy balance, it is rugged and portable. It can also be used as a teaching demonstration for illustrating some of the properties of magnetic circuits.

In these four devices the magnetic field is created by a permanent magnet. It is also possible to use the magnetic field associated with an electric current, and this leads to several more applications. The simplest is the d.c. clip-on ammeter, which measures the current in a wire in terms of the magnetic field developed (Ross and Saker 1955). A small magnetic core is arranged around the wire, and the Hall unit is inserted in an air gap in the core. Currents down to 100 mA can be measured without amplification, with the drive to the Hall unit supplied by a local battery. Smaller currents can be measured if the unit is used in conjunction with the transistor oscillator–amplifier combination, and this technique has been employed for the monitoring of ion and electron beam currents (Fig. 6). The system is useful for currents down to 5 μA (Whitlock and Hilsum 1960).

A group of applications can be envisaged in which the Hall voltage is the result of oscillating currents and magnetic fields. The current may be taken as $I\cos\omega_1 t$ and the field as $H\cos(\omega_2 t + \phi)$, ϕ being a constant. If ω_1 is equal to

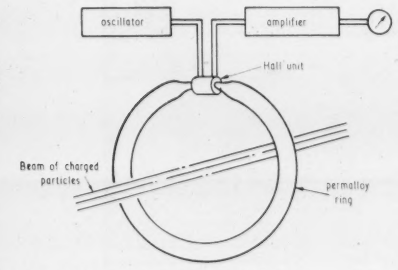


Fig. 6. The measurement of ion and electron currents by means of the Hall effect.

ω_2, ϕ represents the phase difference and the Hall voltage contains a d.c. term, $IH\cos\phi$. This property is employed in the frequency spectrum analyser (Bogomolov 1956 a). Here the current drive is the sum of terms like $I_m\cos\omega_m t$, and the various components are isolated by sweeping the frequency of the magnetic field through the range to be analysed. Wherever the magnetic field frequency is the same as the frequency of one of the components of the current drive, a d.c. Hall voltage is developed.

In the general case the Hall voltage is accurately proportional to the product of the current and the magnetic field. The Hall effect can therefore be used for a simple analogue multiplier, and much of the interest in galvanomagnetic devices has come through this particular application. Several descriptions of Hall effect multipliers have already appeared (Chasmar and Cohen 1958, Hartel 1954, Hilsum 1958) and we need only discuss here the merits of various semiconductors. In a multiplier which is working into a matched load InAs is markedly superior to other materials. InSb would give a higher efficiency, but the accuracy would be low at high input voltages because the unit might suffer appreciable rises in temperature during operation. Further, the matching of the unit to the load would be upset if H were large, since InSb shows a high magnetoresistance effect. The choice of material is less obvious for a multiplier which is to be followed by an amplifier. The factor of merit, $R_H/\sqrt{\rho}$, indicates that germanium would be preferable to InAs, and Lofgren (1958) has deduced from a detailed analysis that even lower mobility materials, such as silicon, may be better still. It is likely that GaAs may become more widely used for this form of multiplier when samples are freely available.

Some other applications, which are similar in principle, are wattmeters for electric power lines (Strutt 1959), modulators for mixers or d.c. amplifiers (Bogomolov 1956 b), square-law detectors (Bogomolov 1956 a), and amplifiers (Ross and Thompson 1955). Barlow (1955) has developed a special form of 'wattmeter' for measuring the power of microwaves in a waveguide. The Hall unit can be arranged in the guide so that the electric vector is parallel to the surface and the magnetic field perpendicular to it. The Hall effect amplifier is a low-noise low-frequency amplifier, but it has not yet found a practical use. In this application the input signal is applied to the magnet coils and develops a magnetic field (Fig. 7). A local battery supplies power to

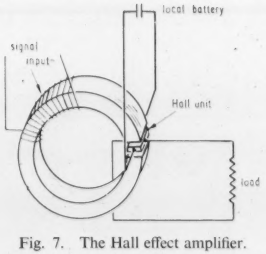


Fig. 7. The Hall effect amplifier.

the Hall unit, but none of the power passes to the load, which is across the Hall electrodes, until the magnetic field is created. The power diverted to the load may be larger than the power needed for the field coils if a high mobility semiconductor is used. InSb Hall amplifiers have shown a power gain of 5 at room temperature (Ross and Thompson 1955). If the

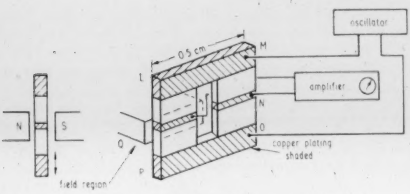


Fig. 8. The magnetoresistance transducer.

output from the Hall unit is fed back to the input coils, the system may be used as an oscillator. However, Bogomolov (1957) has shown that the efficiency is always low.

7. Applications of magnetoresistance

All of the Hall effect applications described above involve a small magnetic field and, following our criterion given earlier, we would not expect them to be suitable applications for magnetoresistance effects. But in some cases it is possible to apply a fixed magnetic field of several thousand oersteds to the sample, and then, though the differential change may be only a few oersteds, the magnetoresistance effect can be used with advantage. For instance the magnetoresistance amplifier is similar to the Hall amplifier shown in Fig. 7, but an additional biasing coil is wound on to the core, and the semiconductor sample is shaped to give large magnetoresistance effects (Thuy 1954). Such an amplifier could in theory give a higher power gain than the Hall amplifier. The magnetoresistance oscillator is also superior to the Hall oscillator and it can give efficiencies up to 37% (Bogomolov 1957). Magnetoresistive current regulators and variable resistances have found some practical use (Willardson and Beer 1956), but the application which will probably be the most successful is the displacement transducer (Ross and Saker 1957). In this device a small permanent magnet provides a field of about 10000 oersteds in a narrow air gap. A sample of InSb or InAs of the shape shown in Fig. 8 acts as the four resistances of a Wheatstone bridge, and the sample is constructed so that the bridge is balanced when it is not in a magnetic field. If the arm LQP is placed in the air gap of the magnet, both LQ and QP increase in resistance, and in one position, with Q at the middle of the air gap, the bridge is again in balance. The balance can now be destroyed by a vertical movement of the magnet, since one arm of the bridge will increase in resistance and another will decrease. The voltage output from the bridge is proportional to the displacement, provided the movement is not large enough to bring the end contacts within the air gap. In a typical InSb

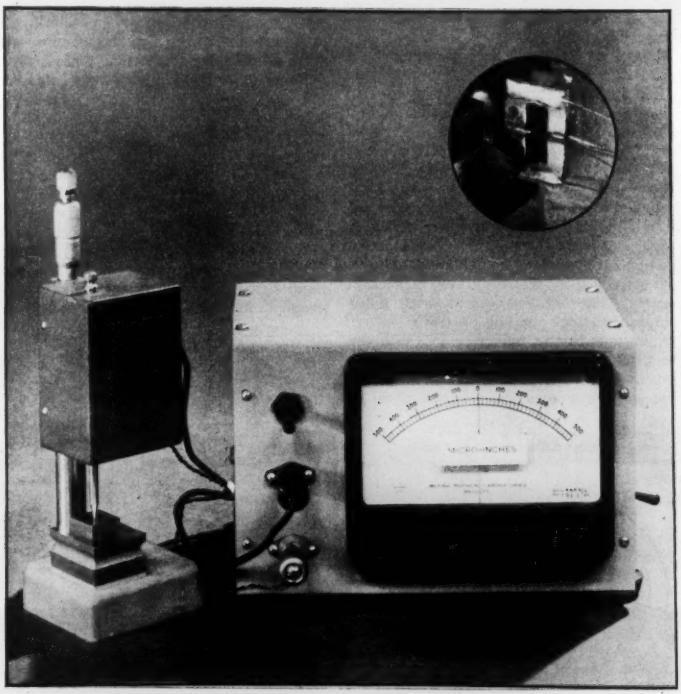


Fig. 9. Displacement transducer. Inset shows magnified view of the InSb element and the magnet.

transducer the sensitivity is 1 volt per inch for displacements up to 0·02 inch. The minimum displacement which should be detectable is near 0·05 Å, if it is assumed that the sensitivity is limited by Johnson noise in a bandwidth of 1 c/s, but it has been found in practice that thermal drifts prevent the detection of movements less than 5 Å (Hilsum and Knight 1959, unpublished). A simple instrument for workshop use (Fig. 9) has the current drive taken from the a.c. mains via a step-down transformer, and the output amplified by a transistor, rectified, and displayed on a milliammeter. Full-scale deflection on the meter corresponds to a displacement of 0·0005 in. (Hilsum and Thompson 1959, unpublished).

The same principle can be used for a microphone or gramophone pick-up. The preferred semiconductor for magnetoresistive applications is p-type InSb, but if the highest sensitivity is not required InAs can be used.

8. Conclusion

As yet none of the devices described above has found wide use in industry. Few complete instruments have been brought forward by commercial manufacturers, and there has been little awareness of the potentialities of this field. But a range of instruments is now becoming available. Multipliers, clip-on ammeters, susceptibility meters and magnetoresistance transducers may soon be common items of equipment in the laboratory and factory. And as greater use is made of these developments, so new devices will be invented and better semiconductors discovered. Exploitation of the galvanomagnetic effects is beginning, but the full possibilities will not be seen for many years.

Acknowledgment

The author is grateful to the Admiralty for permission to publish.

References

Barlow, H., 1955, *Proc. Instn Elect. Engrs*, Part B, **102**, 179.
Bogomolov, V. N., 1956a, *J. Tech. Phys. U.S.S.R.*, **26**, 693.
—— 1956b, *Soviet Physics-Tech. Phys.*, **1**, 2397.
—— 1957, *Soviet Physics-Tech. Phys.*, **2**, 594.
Chasmar, R. P., and Cohen, E., 1958, *Electronic Engng*, **30**, 661.
Elpatevskaia, O. D., and Regel, A. I., 1956, *Soviet Physics-Tech. Phys.*, **1**, 2390.
Hartel, W., 1954, *Siemens Z.*, **28**, 376.
Hilsum, C., 1958, *Electronic Engng*, **30**, 664.
—— 1960, *Solid State Physics in Electronics and Telecommunications*, **2**, 733.
Hilsum, C., and Rose-Innes, A. C., 1958, *Nature* (*London*), **182**, 1082.
Lofgren, L., 1958, *J. Appl. Phys.*, **29**, 158.
Oliver, D. J., 1957, *Proc. Phys. Soc.* B, **70**, 331.
Ross, I. M., and Saker, E. W., 1955, *J. Electronics*, **1**, 223.
—— 1957, *Nature* (*London*), **179**, 146.
Ross, I. M., Saker, E. W., and Thompson, N. A. C., 1957, *J. Sci. Instrum.*, **34**, 479.
Ross, I. M., and Thompson, N. A. C., 1955, *Nature* (*London*), **175**, 518.
Strutt, M. J. O., 1959, *Electronic and Radio Engr*, **36**, 2.
Thuy, H., 1954, *Arch. Elekt.* **8**, 269.
Voeikov, D. D., 1958, *J. Tech. Phys. U.S.S.R.*, **28**, 2248.
Weiss, H., 1960, *Solid State Electronics*, **1**, 225.
Welker, H., 1955, *Elektrotech. Z.*, **76**, 513.
Whitlock, W. H., and Hilsum, C., 1960, *Nature* (*London*), **185**, 302.
Willardson, R. K., and Beer, A. C., 1956, *Electrical Manufacturing*, January.

The preparation of thin films of germanium and silicon

by B. A. IRVING, M.A., Associated Electrical Industries, Aldermaston, Berks.

MS. received 4th November 1960

Abstract

Films about 1000 Å *thick, suitable for observing dislocations by transmission electron microscopy, have been prepared from the brittle semiconductors germanium and silicon by a combination of mechanical polishing and chemical etching.*

Polishing jig

CONTROLLED grinding and polishing was carried out with the jig (Fig. 1), essentially a 1 in. diameter brass block threaded 40 turns per inch and fitted with two steel guard rings. The faces of the block and the guard rings were perpendicular to the axis.

The jig was prepared by locking the rings so that one surface was flush with the face of the brass block and the whole was ground and polished. A locating scratch was then made on the polished face across the guard ring into

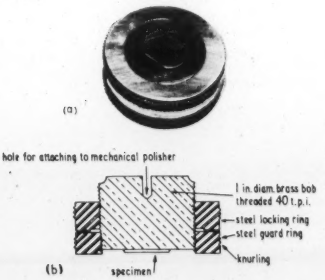


Fig. 1. Polishing jig.

the brass. The rings could then be freed and re-locked so that they protruded a known amount beyond the level of the brass.

For grinding the specimen the same jig was used but both rings were inverted so that what was the upper face of the locking ring protruded beyond the face of the brass and acted as a guard ring. This preserved the polished face of the guard ring with its locating scratch until the final polishing stages.

Grinding was carried out by hand using moistened 600 grade carborundum paper. Diamond dust (final grade 1 micron) was used on a mechanical polisher to produce metallurgically polished surfaces.

Specimens were fixed to the jig with deKhotinsky cement. They were freed by dissolving the cement in hot, peroxide-free, 2 methoxy- or 2 butoxy-ethanol (methyl or butyl cellosolve).

Etching

The polished surfaces were cleaned just prior to etching by immersion in a solution of potassium permanganate in concentrated sulphuric acid (0·01% weight/volume) followed by thorough rinsing with distilled water.

The etchant used for (111) surfaces of germanium was warm (30–40° C) dilute sodium hypochlorite solution (~1% w/v available chlorine). It had been shown previously that this produced a smooth surface, featureless in the electron microscope. Dilute hydrogen peroxide admixed with hydrofluoric acid and alkaline potassium ferricyanide are suitable etchants for (110) and (100) surfaces respectively (Holmes 1959), but have not been used in this application.

For silicon the etchant consisted of a 4% w/v solution of sodium hydroxide to which sodium hypochlorite solution (10% w/v available chlorine) was added dropwise until hydrogen evolution with silicon was just suppressed. It has been used for both (111) and (100) slices.

Scratches, either intentional or resulting from poor grinding and polishing, were enlarged by etching to form shallow grooves. These were generally undesirable but use was made of them to estimate the time required to reduce the thickness of films by etching.

Procedure

(*a*) *First side.* A slice, about 1 mm thick and parallel to a (111) plane, was cut from a germanium single crystal and cemented to the jig. The exposed surface was ground and polished as described above. The slice was then removed from the jig and cleaned. It was placed in a small beaker with the prepared side uppermost and the etchant poured over it and left for about 1 hour. The layer of material disturbed by the grinding and polishing was thereby removed.

(*b*) *Thinning to about* 5 *microns.* The specimen was mounted on the jig with the first prepared side against the polished face of the brass block, and thinned by grinding in stages to about 0·05 mm. The jig was readjusted for polishing and the thickness reduced on a mechanical polisher with 3–5 μ diamond dust to about 10–15 microns. A final readjustment of the jig so that the guard rings protruded about 5 μ and polishing with fine diamond compound (1 μ) thinned the specimen as far as mechanical means safely permitted.

(*c*) *Etching to about* 1000 Å. The specimen was removed from the jig, cleaned and etched by floating on the etchant with the first prepared side uppermost until thin enough to transmit an electron beam in the microscope. This required several hours. Control of the thickness was achieved by simultaneously dissolving a waste piece of the slice in the same etchant. An intentional light scratch made on the

first side after polishing was enlarged to a shallow groove by etching. As the second side was etched away and the thickness was reduced to the depth of the groove, the slice split in two. The waste piece was then immersed completely in the etchant so that it was attacked from both sides. The time taken for this piece to disappear completely was measured and the etching of the floating specimen stopped several minutes before a further similar time interval had elapsed.

Adhesion of the thin films to the electron microscope grids was improved by coating the latter with polyisobutylene. They were dipped into a dilute (2%) solution in chloroform and allowed to dry.

The specimen could be removed from the grids with chloroform, cleaned and etched further if required.

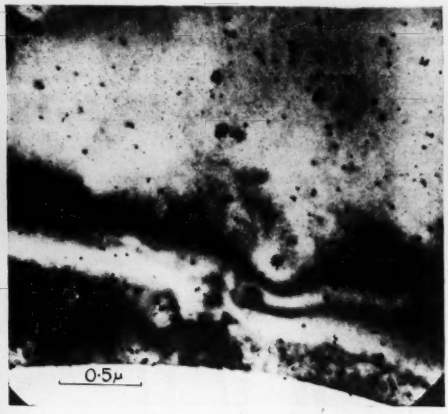


Fig. 2. Electron micrograph of as-grown germanium.

Preliminary reports of the results obtained with specimens prepared by this technique have already appeared (Geach, Irving and Phillips 1957, Phillips and Hartree 1960). Two micrographs are reproduced here of specimens prepared by the above method (Figs 2 and 3). They show extinction contours near the edge of the specimen and minute particles of foreign matter. Fig. 2 demonstrates that the preparative technique eliminates dislocations produced by grinding and polishing. Fig. 3 shows that an array of dislocations in the bulk of the material is not selectively etched out but can be observed in the electron microscope.

No attempt has been made to use this technique to prepare large area specimens; the only limits would be set by the physical size of the grinding and polishing jig and the fragility of the specimens themselves. No absolute measurements of specimen thickness have been required for the present purposes, but thicknesses have been estimated from

Fig. 3. Electron micrograph showing dislocations introduced by indenting at 800° c. Indented surface ground flat prior to preparing thin film.

the transparency of the specimens to an electron beam, and from the approximate determinations of the etch rate implicit in the final etching technique (the thickness after polishing is known from the setting of the polishing jig). This latter method could be used to prepare specimens of intermediate thickness.

Acknowledgments

The author wishes to thank Dr. G. A. Geach for his encouragement and interest in the work and Dr. T. E. Allibone, C.B.E., F.R.S., Director of the Laboratory, for permission to publish this paper.

References

Geach, G. A., Irving, B. A., and Phillips, R., 1957, *Research*, **10**, 411.

Holmes, P. J., 1959, *Acta Met.*, **7**, 283.

Phillips, R., and Hartree, O. P., 1960, *Brit. J. Appl. Phys.*, **11**, 22.

Capillary viscometers of negligible kinetic energy effect

by W. A. CAW, B.Sc., and R. G. WYLIE, M.Sc., Ph.D., Commonwealth Scientific and Industrial Research Organization, Division of Physics, National Standards Laboratory, Chippendale, N.S.W., Australia

MS. received 20th April 1960, in final form 2nd November 1960

Abstract

By incorporating capillaries with long-flared ends in otherwise conventional U-tube or suspended-level viscometers, it is possible to produce instruments in which no significant kinetic energy effect occurs for the Reynolds numbers generally encountered in viscometry. A detailed experimental study of their behaviour has been made. Using the new viscometers a viscosity scale in absolute units can be established with greater accuracy and greater facility than has hitherto been possible.

1. Introduction

ACCURATE viscosity measurements are generally made on what is, in effect, an international viscosity scale based on an adopted value for the kinematic viscosity of water at 20° C (1·0038 cs) and on the use of glass capillary viscometers (American Society for Testing Materials 1953, British Standards Institution 1957, Deutscher Normenausschuss 1955, Institute of Petroleum 1958). With such viscometers, including those incorporating the suspended-level feature due to Ubbelohde (1937) the observed time of flow t is taken to be related to the kinematic viscosity ν by the equation

$$\nu = Ct - B/t \qquad (1)$$

where C and B are constants for a particular instrument. The second term, or kinetic energy correction, which is by no means negligible for the lowest viscosities normally encountered, has been the subject of much investigation.

A simple theory due to Hagenbach (1860) and Wilberforce (1891) gives for B the explicit expression $V/8\pi l$, where V is the volume of flow and l is the capillary length. However, the theory is at best only approximately correct and it is usual to express B in the form

$$B = mV/8\pi l \qquad (2)$$

and to focus attention on the numerical factor m. However, Knibbs's (1895, 1896, see also Barr 1931) analysis of the experiments of Poiseuille, and the results of Bond (1922) and of Bell (1947), among others, show that m, although often taken to be 1·12, cannot be assumed to be independent of flow time at the Reynolds numbers encountered in viscometry. Nevertheless, the standard method of establishing the viscosity scale treats B as a constant. The kinetic energy correction permitted is about 1% within the British and American specifications, and 2% within the German specification.

As long as significant kinetic energy effects occur in the instruments used to establish the viscosity scale, those effects will contribute fundamental difficulties and complications, and preclude any worthwhile improvement in the accuracy.

A new type of viscometer has been developed in which the kinetic energy effect is much less than in equivalent conventional instruments and is negligible for most purposes. The new viscometers are of about the same size as conventional types and hardly more difficult to construct. Their use in the establishment of the viscosity scale affords an immediate improvement of 0·2% in accuracy. They also open up valuable new methods, in particular fulfilling the long-standing need for a method of detecting very small deviations from Newtonian behaviour in low-viscosity liquids.

2. Design and construction of the viscometers

In the design of the viscometers the concept of a uniform capillary with identifiable ends is abandoned. Instead, the aim is to achieve a hydrodynamical resistance, which is constant up to the highest relevant Reynolds numbers through the absence of any marked curvature of the stream lines in regions that make a significant contribution to the resistance. Such a constant resistance can be provided by a capillary of one of the forms (*b*) and (*c*) shown in Fig. 1,

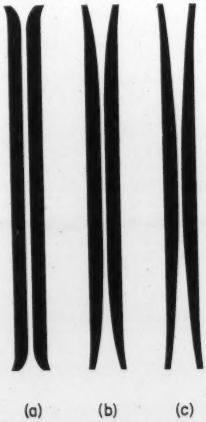


Fig. 1. Conventional and long-flared capillaries. (*a*) conventional; (*b*) long-flared with parallel centre section; (*c*) with overlapping long flares.

in which the diameters are exaggerated. Fig. 1(*a*) represents a conventionally shaped capillary for purposes of comparison. The significant feature in (*b*) and (*c*) is a slow, monotonic increase in diameter and longitudinal curvature of the wall from the commencement of the flare, or the centre of the tube, to the extremity.

A significant kinetic energy effect can exist at flow rates well below those for which jet formation occurs at the exit orifice. The effect is then due more or less equally to inlet and exit orifices. To produce a large reduction in the value of B equal attention must, therefore, be given to each end of the capillary. Experimental proof that the kinetic energy effect can be reduced greatly in practice will be given below. A viscometer embodying a long-flared capillary may still incorporate the suspended-level feature, with the usual advantages.

The required capillary forms can be produced by subjecting a uniform glass capillary tube to pneumatic pressure while it hangs in a tube furnace in which there is a temperature gradient. On the assumption that the viscosity of glass varies exponentially with the inverse temperature, it can be shown that in a uniform temperature gradient a tube of internal diameter a acquires a suitable form such that its variable diameter d is given approximately by

$$d = a\{1 + \exp(x/\beta a)\} \tag{3}$$

where β is a shape factor and where x is measured from the point of diameter $d = 2a$. The physically significant parts of the capillary, as regards the constancy of the overall flow resistance, are those in which the diameter increases from about 1·01 times to about twice its minimum value. These parts can be formed wholly within a uniform temperature gradient.

3. Experimental investigation of the kinetic energy effect

An experimental investigation has been made of the behaviour of viscometers incorporating capillaries flared to different degrees. The results, which extend over three decades of the Reynolds number, show a progressive and marked reduction in the kinetic energy effect with increasing degree of flare.

3.1 *Theoretical aspects.* Although the experiments range over conditions for which Eqn (1) does not hold with C and B constant, that equation can be retained if C and B are suitably defined. One approach is to define the coefficients locally, that is, as having those values for which Eqn (1) fits the viscosity–time curve in the region in question. However, in the following, except where the designation 'local' is used, C is understood to be a constant, namely, the quotient of a very large viscosity and the corresponding very long flow time; B is then simply that quantity for which Eqn (1) is valid.

It will be shown that, in the regions of particular interest, the experimental results correspond to a relationship between viscosity and time of flow which is of the form

$$\nu = Ct - K/t^n \tag{4}$$

where C and K are constant for a particular viscometer. B, defined by Eqn (1), is then of the form

$$B = K/(t^{n-1}). \tag{5}$$

It can be shown that Eqn (4) implies the more specific equation

$$\nu = M\frac{h}{V}t - \frac{NV^n}{h^{\frac{1}{2}(n-1)}t^n} \tag{6}$$

where M and N are constants associated with a particular capillary and h is the effective pressure head. In practice, the pressure head changes during flow by a considerable fraction, which depends on the design of the particular instrument. Strictly speaking, therefore, an equation of flow that is determined experimentally for one instrument is of direct significance only for instruments of the same design. However, for present purposes the variation of the head during flow is of little consequence; it can be shown that for the values of n necessary to fit the experimental results of this paper, if Eqn (6) is valid at constant head with particular values of n and N, then, with those same values of n and N and an only slightly modified value of M, it is applicable with good accuracy when the head changes during flow by as much as 2 : 1, a mean value being adopted for h. Thus, where Eqn (4) fits the experimental results, the values of n and K derived from those results are not restricted in significance to flow with a changing head, but may be regarded as applying to the ideal case of flow at a constant head, namely the mean head.

To obtain accurate information concerning the correction term, specially designed U-tube viscometers of the form shown in Fig. 2 were used. The fall of the meniscus may be timed between the marks a and b and also between b and c.

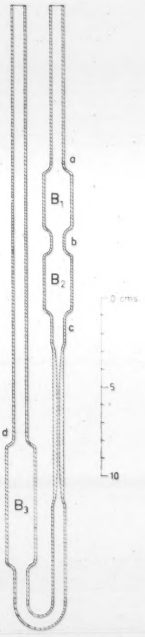


Fig. 2. The form of viscometer used in the investigation. a, b and c, timing marks; d, filling mark; B_1 and B_2, upper and lower timing bulbs; B_3, reservoir bulb.

The instrument is, therefore, equivalent to two viscometers with a common capillary. If t_1 and t_2 are respectively the times for flow from the upper and lower bulbs, and if α is the corresponding ratio $V_2{}^n h_1{}^{\frac{1}{2}(n-1)}/V_1{}^n h_2{}^{\frac{1}{2}(n-1)}$, then the viscosity may be expressed both as

$$\nu = C_1 t_1 - \frac{K}{t_1{}^n} \text{ and } \nu = C_2 t_2 - \frac{\alpha K}{t_2{}^n}. \tag{7}$$

Eliminating ν in Eqns (7), one obtains

$$\frac{C_1 t_1}{C_2 t_2} - 1 = K\left(\frac{t_2^n/t_1^n - \alpha}{C_2 t_2^{n+1}}\right). \qquad (8)$$

In the instruments used, the volumes V_1 and V_2 have been made equal to within 2%. Consequently, for those with conventional capillaries, for which n is approximately unity over a wide range of the Reynolds number, the value of α is approximately unity, whereas for those with long-flared capillaries $\alpha \simeq (h_1/h_2)^{\frac{1}{2}(n-1)}$ in the range of Reynolds number for which Eqn (4) holds. The appropriate value of n is that for which the quantity on the left-hand side of Eqn (8), plotted against the quantity in brackets using data representing a series of viscosities, gives a straight line. The value of K is simply the slope of this line.

It can be shown that if Eqn (8) is found to hold with K constant, then Eqn (4) must apply.

3.2 *Experimental.* The results given have been obtained with a series of four viscometers of the type shown in Fig. 2. The lengths of the flares range from conventional (viscometer I) to about half the length of the capillary (viscometer IV).

Definition of shape and size of capillaries for viscometers I–IV

	I	II	III	IV
a (mm)	0·47	0·47	0·49	0·48
β	1·7	5·8	11·2	36
f	69	82	98	220

The capillary of instrument IV is more nearly of the form of Fig. 1(*c*) than of Fig. 1(*b*). The four instruments are designed so that their flow times are much the same, and so that in instrument I the kinetic energy effect provides an unusually large proportion of the total flow resistance (10% at 1 cs for flow from the upper bulb).

Measurements have been made with each instrument using a series of liquids, each at a number of different temperatures. The liquids were, in order of decreasing viscosity, di-butyl phthalate, aniline, dioxane, water, benzene, methyl ethyl ketone and acetone. The concordance of the data obtained with these different liquids in any one instrument, which is apparent from the plots of Figs 3, 4 and 5, shows that all the liquids behaved in a Newtonian manner.

The design of the experiment is such that conclusions

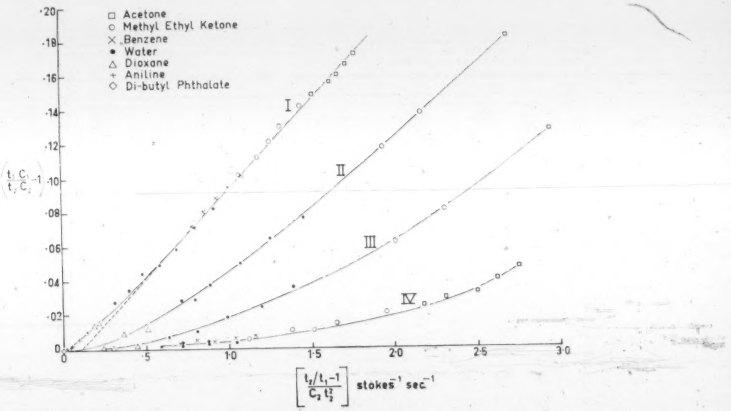


Fig. 3. Plot of the results on the basis of the conventional equation of flow (Eqn (1) of the text). A linear plot indicates that the equation applies.

Whereas the conventional capillary flares were formed by ordinary glass-blowing techniques, the long flares were produced by subjecting each end of the Pyrex capillary tube in turn to a uniform temperature gradient, while applying internal pneumatic pressure. The temperature at the extremity of the tube was about 730° C while the gradient was a few degrees per centimetre for instrument IV and progressively greater for III and II. In each case a pressure of 20 lb/in^2 was applied for about 5 min. The flares of tubes II, III and IV are closely in accordance with Eqn (3), while those of tube I are approximately so. The parameters defining the shape and size of the capillaries are given in the Table. The distance from the origin of x for the flare at one end of the capillary to the origin for the flare at the other end, expressed as a multiple of a (Eqn (3)), is denoted by f. The capillaries range in overall length from approximately 3·5 cm for instrument I to approximately 12 cm for instrument IV.

drawn after fitting the results to Eqn (8) are largely unaffected by most of the types of error that arise in conventional viscometry. Thus the temperature need not be known, provided it is the same for flow from the upper and lower bulbs. This condition was satisfied to $\pm 0{\cdot}002$ deg C, the corresponding error being quite negligible. Also, a single fill of liquid provides flow from both the upper and lower bulbs. The method is very insensitive to any constant error in the rate of the timing device. If the liquids have the same surface tension, no surface-tension error of any consequence can occur. The surface tensions of all the liquids except water are, in fact, roughly the same, not varying much from 30 erg/cm^2. Although the surface tension of water is much higher than this, being approximately 73 erg/cm^2, the points for water in the plotted results show no systematic deviation as compared with those for the other liquids, confirming that surface-tension forces largely balance out. The volume of liquid retained on the wall of a bulb during flow can be

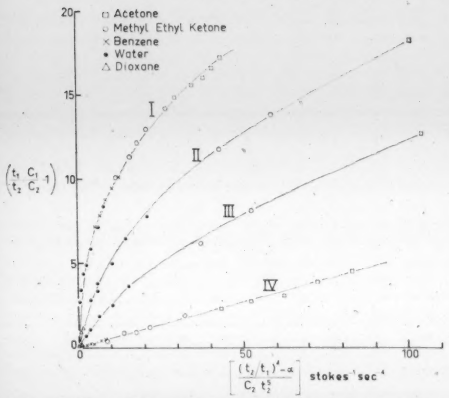


Fig. 4. Plot of the results on the basis of the equation $\nu = Ct - K/t^4$. A linear plot indicates that this equation applies.

shown to be independent of viscosity at constant surface tension; further, it can be shown to be very insensitive to surface tension and was of no consequence in the present experiments. However, care had to be taken to run the liquid back sufficiently slowly between measurements to leave only an insignificant amount of liquid on the wall of the reservoir bulb. A reproducibility of the vertical alignment of the instruments to 0·2° or better is important, but was easily achieved in the experiments. Likewise care was necessary to fill the instruments accurately to the appropriate mark. In fact, the only significant source of error appears to be the observation of the times of flow, which gives rise to an unbiased scatter of the results.

3.3 *Results.* The results are most readily interpreted when plotted on the basis of Eqn (8), as in Fig. 3 with $n = 1$ and in Fig. 4 with $n = 4$. It is immediately apparent from these plots that the kinetic energy effect diminishes progressively and markedly from viscometer I to viscometer IV.

In Fig. 3 the plot for the instrument with conventional flares (I) is closely linear, except at low flow rates. The disposition of the points in this plot shows evidence of significant small deviations from the straight line that do not appear to be internal to groups of points corresponding to the same liquid. Whether real or not, these deviations have no bearing on the present discussion. The plots for instruments II and III are non-linear at low flow rates but apparently tend to straight lines at high flow rates. Thus, certainly for instrument I and probably also for instruments II and III, with increasing flow rate the results tend to conform to Eqn (1) with constant *local* values of C and B. The linear part of the plot for I corresponds to a value of m of $1{\cdot}25 \pm 0{\cdot}05$. This approximate constancy of m with a conventional capillary from a Reynolds number of about 200 up to at least 600 (defined as $4Q/\pi\nu a$, where Q is the volume rate of flow for the upper bulb) is in contrast with the results of Bell (1947), who reported that m varied from about 0·6 to 0·8 over that range.

The plot for instrument IV in Fig. 3 is curved up to the highest flow rate used, corresponding, in terms of flow from the upper bulb, to a Reynolds number of approximately 1600.

In Fig. 4 the plot for instrument IV is a straight line over the whole range. The behaviour of this instrument thus accords with Eqn (8) with $n = 4$. Indeed, if the true value of n had differed from 4 by even 10%, the plot would have been conspicuously curved. The curves for the other instruments deviate progressively further from the line for instrument IV. The results suggest that the curves for instruments II and III tend to linearity in the neighbourhood of the origin, and that each of these curves approaches the origin at a definite angle to the co-ordinate axes. It is thought that this is the case and, consequently, that Eqn (8), with $n = 4$, is applicable to these instruments at limiting small flow rates.

To focus attention on the region of greatest interest, the data obtained with instruments I and IV at the lower flow rates, including a number of points omitted from Figs 3 and 4 for the sake of clarity, are plotted against the Reynolds number in Fig. 5. It may be noted that in this region the Reynolds number is approximately proportional to the

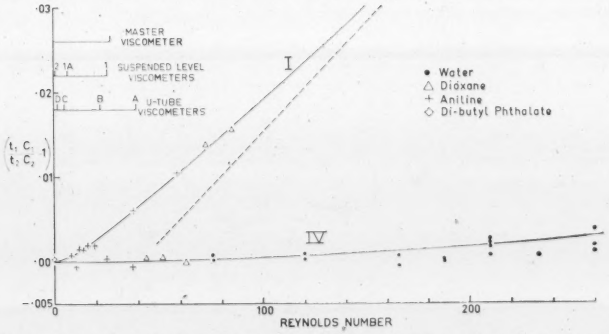


Fig. 5. Plot of the data for the region relevant to standard viscometry with Reynolds number as abscissa. The maximum Reynolds numbers permissible (e.g. according to Brit. Stand. 188: 1957) in a master viscometer and some standard U-tube and suspended-level viscometers are shown.

(1) Master viscometer (Brit. Stand. 1957, Inst. Pet. 1958,) U-tube and suspended level.
(2) Suspended-level viscometers (Brit. Stand. 1957, Inst. Pet. 1958).
(3) U-tube viscometers (Brit. Stand. 1957).

abscissa quantity of Fig. 3, and that the curve drawn for instrument IV in Fig. 5 is equivalent to the straight line in Fig. 4. As in Fig. 3, the broken line represents an extension of the linear part of the plot for instrument I.

Figure 5 also shows the Reynolds numbers corresponding to the most rapid flow that the English and American specifications (American Society for Testing Materials 1953, British Standards Institution 1957, Institute of Petroleum 1958) allow in suspended-level viscometers 1, 1A and 2, U-tubes A, B, C and D, and in master viscometers. For instrument I the local value of B does not become constant until the Reynolds number exceeds the highest value allowed. Thus the use of a constant B-value, as the specifications require, can be no more than a compromise. However, with instrument IV the kinetic energy effect is not detectable for Reynolds numbers up to about four times the highest that these specifications allow. The actual magnitude of the effect in this instrument is now considered.

The slope of the curve for instrument IV in Fig. 4 gives for K the value 540 stokes s^4. The value of α is 2·23. For flow from the lower bulb the principal constant of the instrument is $5{\cdot}49 \times 10^{-5}$ stokes s^{-1}, corresponding to a flow time for water at 20° C of 183·7 s. At this flow time the ratio of the correction term to the principal term is only $1{\cdot}0 \times 10^{-4}$. The corresponding ratio for a British Standard U-tube A, when used with water at 20° C (flow time approximately 300 s), is approximately 5×10^{-3}. Thus, even though the flow time in instrument IV is little more than half that in the B.S. U-tube, and the dimensions of the two instruments are comparable, the kinetic energy correction for the former is 50 times smaller than for the latter. If instrument IV is compared with the standard suspended-level viscometer No. 1 at the lowest viscosity allowed, namely 0·02 stokes, the calculated kinetic energy corrections are found to be in a ratio of almost 2000 : 1.

Clearly, viscometers can easily be constructed with which accurate measurements can be made without kinetic energy corrections, even well below the range of viscosities covered by the present specifications.

A number of general conclusions can be drawn from the curves of Figs. 3 and 4 as to how the flare length and flow rate affect the behaviour of instruments with capillary flares of the present general shape.

(i) At high flow rates: (*a*) except, perhaps, with the longest flares, as the Reynolds number increases the behaviour of an instrument tends to conformity with Eqn (1), with local constants C and B; (*b*) the Reynolds number at which the behaviour conforms to Eqn (1) is higher the longer the flares.

(ii) At low flow rates: (*a*) as the Reynolds number decreases the behaviour tends to conformity with Eqn (4), the value of n being 4; (*b*) conformity with Eqn (4) extends to higher Reynolds numbers the longer the flares; (*c*) the value of K in the equation decreases with increasing flare length.

It would not be surprising to find that the validity of Eqn (4) at sufficiently low Reynolds numbers, with $n = 4$, extends to instruments with capillary ends of any shape, provided that the longitudinal curvature of the capillary wall, in the critical zone, increases monotonically and that, in this zone, the angle between the wall and the capillary axis is small. The latter proviso is not satisfied in instrument I of the present set, and the curved extremity of the plot for this instrument in Fig. 3 probably does not represent an approach to $n = 4$.

In other experiments, long-flared capillaries were incorporated satisfactorily in viscometers of the suspended-level type to give results generally in accordance with those given above. The meniscus representing the suspended level forms readily at the extremity of the long exit flare. The shape of this extremity must, however, be chosen with the usual discretion.

4. Conclusion

It has been shown possible to construct glass capillary viscometers, incorporating long-flared capillaries, in which the kinetic energy effect is negligible for most purposes. A study of the behaviour of a series of instruments with capillary flares, ranging in length from conventional to about half the length of the capillary, points to the conclusion that, except perhaps for the shortest flares, the correction term in the equation of flow tends to the form K/t^4 at low Reynolds numbers.

The manner in which the length of the flares affects the value of K and the range of validity of the inverse fourth-power law is the subject of a further investigation in progress. In other work, viscometers of the new type are being used to determine accurately the relative viscosity of water at a series of temperatures and the effect of dissolved air. The results of these investigations and a more detailed account of the design and construction of the viscometers will be published in due course.

Acknowledgments

The authors wish to thank Dr. G. H. Briggs and Mr. A. F. A. Harper for their helpful criticism, Mr. G. E. Nightingale for his very able construction of the viscometers used, and Mr. W. Nikolsky for the numerous careful observations made.

References

American Society for Testing Materials, 1953, *ASTM D*445–53 T.

Barr, G., 1931, *A Monograph of Viscometry* (Oxford: University Press).

Bell, J. D., 1947, *Thesis* (Pennsylvania State College, School of Chemistry and Physics).

Bond, W. N., 1922, *Proc. Phys. Soc.*, **34**, 139.

British Standards Institution, 1957, *Brit. Stand.* 188: 1957.

Deutscher Normenausschuss, 1955, *German Standard DIN* 51562.

Hagenbach, E., 1860, *Pogg. Ann.*, **109**, 385.

Institute of Petroleum, 1958, *Standard Methods for Testing Petroleum and its Products*, 648 (IP 71/58).

Knibbs, G. H., 1895, *J. Roy. Soc. New S. Wales*, **29**, 77.

—— 1896, *J. Roy. Soc. New S. Wales*, **30**, 190.

Ubbelohde, A. R., 1937, *J. Inst. Petrol. Tech.*, **23**, 427.

Wilberforce, I. R., 1891, *Phil. Mag.*, **31**, 407.

Internal friction in fibre assemblies

by J. D. HUFFINGTON, British Jute Trade Research Association, Dundee

MS. received 25*th July* 1960, *in revised form* 13*th October* 1960

Abstract

Experiments have been carried out to measure the energy dissipated as internal friction (hysteresis loss or work of deformation) when sliding rectangular steel sliders on various types of fibre assembly. The theoretical assumptions previously made by Greenwood and Tabor in explaining experiments with steel balls sliding on rubber lead to predictions of internal friction in fibre assemblies which are too low. It is thought that one reason for this is the tendency for the pressure to build up at the front edge of the slider when it moves over a fibre assembly, thus increasing the amount of internal friction corresponding to a given deformation. It has also been found that the energy dissipated appears to be abnormally high for staple yarns as compared with continuous filament yarns. It is suggested that two mechanisms of energy dissipation exist. On the one hand there is the normal mechanism where deformation is subject to complete elastic recovery, although energy is lost in the process. In addition, with staple yarns in particular, work of internal rearrangement may occur which is not subject to elastic recovery, but which represents a permanent change in the yarn. It is suggested that this may be a wear or abrasion factor due to general sliding and loosening of fibres within the yarn. A quantitative measure of this factor offers an approach to a definition of abradability or liability to wear.

Introduction

In view of its possible technological importance in the textile field it is surprising that no experimental work, as far as the author is aware, has been done to measure the effect of internal friction (hysteresis loss or work of deformation) on the friction of fibre assemblies. The work described here is of a preliminary nature only, but it has been sufficient to show the significance of this phenomenon.

The previous theoretical work on rubber (Greenwood and Tabor 1958) has been helpful in interpreting the results, but these have shown that internal friction in fibre assemblies is different in character from that in rubber. A fibre assembly such as a yarn is a comparatively loose structure with little adhesion between individual fibres. This leads to marked deviations from theoretical predictions based on a rubber-like model, the internal friction in fibre assemblies usually being higher than would be expected on this model. This increase has been most clearly apparent with staple yarn or cloth, where fibre displacements are not necessarily reversible and may lead to a permanent change in the assembly structure. It may be that a quantitative measure of this factor (which is a basic property of the assembly) would give a new approach to the problem of abrasion or wear.

Experimental

The apparatus used was similar to that described previously (Huffington and Stout 1960). Upper rectangular steel sliders were used, about 1 mm thick and 4 cm wide and of different lengths l in the direction of sliding. In most of the experiments the front and rear edges of the slider (those perpendicular to the direction of sliding) were cut at an angle of 70° with the horizontal as in Fig. 5. The lower surface consisted either of a single layer of parallel yarns wound on to a wooden platform in the direction of sliding, or a strip of cloth or felt. The bulk of the results were with yarns, which were spaced sufficiently far apart in the layer that they did not touch when flattened by the slider, thus eliminating any restriction of deformation due to mutual interference.

If the frictional force per unit slider area F/A is plotted against the normal load per unit area $W/A = p$ for steel rectangles of different l values a series of curves as shown by the typical result in Fig. 1 are obtained. As the l value is

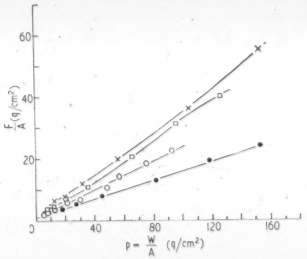


Fig. 1. F/A against $W/A = p$ (g/cm²). Layer of viscose staple yarn 1 mm diam., 3·7 T.P.I., 50 denier.
× $l = 0{\cdot}87$ cm, □ $l = 1{\cdot}4$ cm, ○ $l = 2{\cdot}35$ cm, • $l = \infty$ cm.

increased the slopes of the curves are reduced. This is because the frictional force consists of two terms, the friction between the flat steel surface and the assembly surface (F_s per cm²) and an excess friction located at the front edge (F_e per cm of front edge of width b).

i.e.
$$F = F_s A + F_e b \tag{1}$$

$$\frac{F}{A} = F_s + \frac{F_e}{l}. \tag{2}$$

Thus at large l, or in the absence of excess friction $F/A = F_s$. The (F/A, p) curve in Fig. 1 corresponding to this condition was obtained by sliding the assembly over a steel lower surface of larger area in such a way that excess friction at the front could not occur. The assembly and the steel were identical with those used in the other experiments referred to above. The lowest curve in Fig. 1 shows the result obtained. From Eqn (2)

$$\log\left(\frac{F_e}{l}\right) = \log\left(\frac{F}{A} - F_s\right). \tag{3}$$

In Fig. 2 $\log (F/A - F_s)$ is plotted against $\log p$ for different l values, using the experimental results of Fig. 1. These curves are displaced, roughly parallel to one another, depending on the value of l. If the average value of $(F/A - F_s)$ over the experimental range of pressures is obtained at each l from

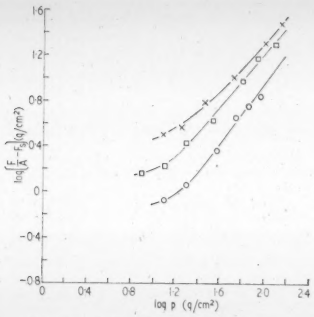


Fig. 2. $\log (F/A - F_s)$ against $\log p$ (g/cm^2). Description of points as in Fig. 1.

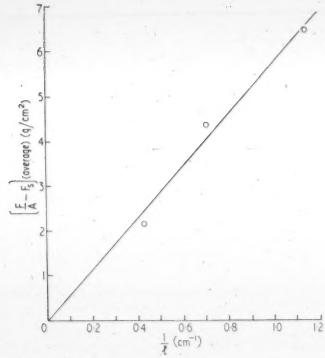


Fig. 3. $(F/A - F_s)$ (average) (g/cm^2) against $1/l$ (cm^{-1}) from Fig. 2.

Fig. 2 and plotted against $1/l$, Fig. 3 is obtained. Since this is a straight line passing through the origin it shows that the results are consistent with Eqn (2). Also Fig. 2 shows that $F_e = (F/A - F_s)l$ is roughly proportional to p except at low pressures.

In Fig. 4 F_e is shown as a function of p obtained from the experimental results referred to above for $l = 1{\cdot}4$ cm. Also shown is the result obtained under similar conditions with a layer of nylon continuous filament yarn. Results similar to those described above have been obtained for a number of different fibre assemblies, and it is of interest to consider how far they can be understood theoretically.

The excess friction at the front edge
Theoretical considerations

Suppose the front edge of a rectangular steel slider moves perpendicularly along the axes of a number of parallel yarns and makes an angle β with the horizontal as in the vertical sectional view in Fig. 5 (bottom). Viewed from above the region of contact between slider and a single yarn will be as in Fig. 5 (top).

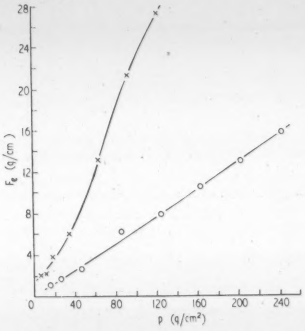


Fig. 4. F_e (g/cm) against p (g/cm^2).
× Layer of viscose staple yarn 1 mm diam., 3·7 T.P.I., 50 denier.
○ Layer of nylon continuous filament yarn 1 mm diam. as in Table 1.

Let p_e be the pressure during sliding normal to the interface averaged over the whole area of the slider edge. (The p

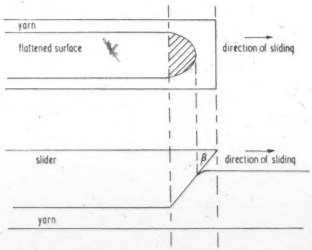


Fig. 5. The contact region at the front edge for a single yarn (β = constant).

previously referred to was the value averaged over the whole area of the flat part of the slider.) Then the corresponding load sustained at the edge is given by

$$W = p_e \cos\beta - \mu p_e \sin\beta \qquad (4)$$

where μ is an average value of the interfacial coefficient of friction. Similarly the corresponding force component in the direction of sliding is

$$F = p_e \sin\beta + \mu p_e \cos\beta. \qquad (5)$$

The results for the interfacial coefficient of friction of steel against the different materials considered in this paper (without excess friction at the edge) did not exceed 0·2. Eqns (4) and (5) become, with $\beta = 70^\circ$ and $\mu = 0{\cdot}2$

$$W = 0{\cdot}34p_e - 0{\cdot}19p_e \qquad (6)$$

$$F = 0{\cdot}94p_e + 0{\cdot}07p_e. \qquad (7)$$

Thus the contribution of interfacial edge friction to F is

comparatively small, although it appreciably influences W. Eqn (6), however, does not enter into the calculation of the results.

From (5), the force per unit width of slider (neglecting interfacial friction) is

$$F = p_e t \tag{8}$$

where t is the total thickness of the slider.

If we assume, as a maximum estimate, 100% hysteresis loss, the excess frictional force per unit width of leading edge becomes

$$F_e = p_e t. \tag{9}$$

The pressure distribution over the leading edge and the area of contact (shown shaded in Fig. 5 (upper)) are unknown, although visual observation showed that the thickness of slider involved in contact with the yarn was less than t. Again as a maximum estimate it is provisionally assumed that $p_e = p$. Since $t = 0{\cdot}114$ cm

$$F_e = 0{\cdot}114p = Bp. \tag{10}$$

The experimental results have indicated a variation of F_e with p of roughly the same form as in Eqn (10) in all cases (see, for example, Fig. 4). However, the constant B deviates somewhat from the estimated figure as shown by the results in Table 1. In view of the assumptions made the B values in Table 1 seem reasonable for continuous filament yarn and felt, but all the results for staple yarn and cloth (except one) exceed the maximum possible estimate given above.

Table 1. *Rectangular steel sliders on fibre assemblies* $\beta = 70°$, $t = 0{\cdot}114$ cm

Assembly	B (exptl)	$\frac{B\,(\text{exptl})}{0{\cdot}114} = \gamma$	*Log (No. of rubs) Walker abrasion test*
Layer of nylon cont. filament yarn (1 mm diam., twist 2 or 7 T.P.I.)	0·064 (Fig. 4)	0·56	—
Layer of nylon staple yarn (1 mm diam., 2 T.P.I., 6 denier)	0·14	1·23	2·14
Layer of nylon staple yarn (1 mm diam., 7·2 T.P.I., 6 denier)	0·085	0·75	4·48
Layer of viscose staple yarn (1 mm diam., 3·7 T.P.I., 4·5 denier)	0·15	1·32	2·25
Layer of viscose staple yarn (1 mm diam., 3·7 T.P.I., 50 denier)	0·22 (Fig. 4)	1·93	1·11
Wool felt	0·064	0·56	—
Jute cloth	0·144	1·26	—

It is clear that a factor has been omitted from the theory given above. One assumption implicitly made was that the internal friction was a function only of the deformation (as with rubber), so that static loading experiments with the same deformation would give the same internal friction as in the sliding experiments. It is probable, however, that with a yarn (for example) during sliding there is fibre displacement along the axis of the yarn due to the dragging effect of the front edge. In this case during sliding the yarn no longer has uniformity of structure along its axis, there being a discontinuity at the front edge. This can be allowed for in a crude but simple way by assuming the pressure normal to the interface at the front edge to be γp where γ may be greater than unity. Eqn (10) then becomes

$$F_e = 0{\cdot}114\gamma p \tag{11}$$

$$\gamma = \frac{B\,(\text{exptl})}{0{\cdot}114}. \tag{12}$$

Values of γ in Table 1 suggest that there may be a pressure build up at the front edge with staple yarns and cloth, and that this is considerably less for continuous filament yarns and felt. This is reasonable since, although fibre movement along the axis can occur with continuous filament yarn, it is likely to be more restricted than with staple yarn. Also any fibre displacement in a continuous filament yarn is reversible, since the yarn ultimately recovers its original structure. With a staple yarn this is not so, there being, for example, a noticeable increase in the hairiness of a yarn after an experiment. It thus seems reasonable to correlate the increased value of γ for staple yarns above that for continuous filament yarns with internal sliding and rearrangement of fibres representing a permanent change in the yarn. In other words this may be a measure of the abradability or liability to wear of the assembly.

This idea has been tested by measuring the abrasion resistance of the staple yarns using a Walker yarn abrasion tester (Walker and Olmstead 1945). In the final column of Table 1 the logarithm of the average number of rubs required to break the yarn under a certain tension is shown, and this appears to correlate with the value of γ (the higher γ the lower the abrasion resistance).

The effect of the shape of the front edge of the slider

Only an incomplete investigation has been carried out, but this has been sufficient to show that F_e depends on the shape of the front edge. In one set of experiments the front edge was rounded as in Fig. 6 and this caused a

Fig. 6. Section of front contact edge for a rectangular steel slider with a rounded edge deforming a yarn.

reduction in F_e as compared with the straight edge with $\beta = 70°$. With a rounded edge it is difficult to make a theoretical estimate of F_e, but possibly the effective value of p_e in Eqn (9) can be regarded as less than when $\beta = 70°$. The values of B (exptl) in Table 1 for $\beta = 70°$ are reduced by a roughly constant factor of 0·4 for the rounded edge. Thus F_e values for the different assemblies still have the same ratio to one another, although reduced in magnitude, staple yarns showing once again a relatively high internal friction as compared with continuous filament yarns. The abrasion damage caused with the rounded edge is, of course, less than when $\beta = 70°$.

Relative magnitude of F_s and F_e

The magnitude of the internal friction at the front edge per unit width F_e is perhaps best understood by comparing it with the friction between the surfaces under the slider per unit area F_s. In Table 2 F_e/F_s is shown (averaged over the experimental range of pressures) for different fibre assemblies and for both straight and rounded edges.

Table 2.

Assembly	F_e/F_s ($\beta = 70°$) (cm)	F_e/F_s (rounded) (cm)
Layer of nylon cont. filament yarn, 1 mm diam.	0·48	0·20
Layer of nylon staple yarn, 1 mm diam., 2 T.P.I.	0·91	0·47
Layer of nylon staple yarn, 1 mm diam., 7·2 T.P.I.	0·66	0·32
Layer of viscose staple yarn, 1 mm diam., 3·7 T.P.I., 4·5 denier	0·82	—
Layer of viscose staple yarn, 1 mm diam., 3·7 T.P.I., 50 denier	0·95	—
Wool felt	0·35	0·1
Jute cloth	0·63	0·3

F_e/F_s is by definition the ratio of the two types of friction for a particular dimension (1 cm²) of slider. Enlarging the slider area reduces the relative importance of F_e, and decreasing the area increases its relative importance. For example, for the 50 denier viscose staple yarn and a slider 1 mm square with $\beta = 70°$ the internal friction would be about ten times the surface friction. For continuous filament nylon yarn and a slider 1 mm square with a rounded edge the internal friction would be twice as great as the surface friction. These examples show that the internal friction can be more important than the ordinary friction under suitable experimental conditions.

Variation of F_e with p over an extended range

Figure 4 shows that there is a tendency for F_e to increase more slowly with pressure at both the lowest and highest pressures. This appears to occur fairly generally and Fig. 7

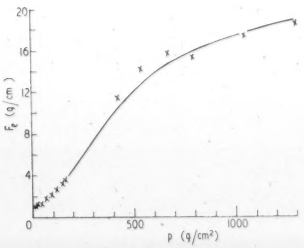


Fig. 7. F_e (g/cm) against p (g/cm²) for rectangular steel slider with a rounded edge sliding over a layer of nylon continuous filament yarns of 1 mm diam.

shows results obtained with nylon continuous filament yarn over a wide range of pressures (slider with a rounded edge). At very low pressures there may be little bulk deformation of the yarn, only a closing of interstices between loose fibres, with a relatively small increase of F_e with pressure. The main variation of F_e with pressure is thought to be due to flattening of the yarn by bulk deformation, but this is restricted at high pressures, when F_e must presumably tend ultimately to a fixed maximum value.

Discussion

The experiments described above have established that the work of deformation when sliding steel rectangles over fibre assemblies is often abnormally high. This is probably because the sliding produces a build up of pressure at the front. For example, when the edge slides perpendicularly along the axis of a yarn there is a temporary increase of twist in front and this means that a given deformation may dissipate more internal friction. With staple yarns there is an additional effect due to the increased freedom of movement of individual fibres with the possibility of the edge displacing them irreversibly, thus altering the structure of the yarn.

The theoretical treatment has been much simplified, but the present knowledge of the deformation characteristics of fibre assemblies is insufficient to allow a more rigorous study. Even if the static deformation properties of the assembly were known, this would not be adequate under dynamic conditions. The static deformation of a yarn when loaded as in Fig. 5 depends mainly on twist, tension of the yarn, elastic properties of the fibre and inter-fibre friction. The results for nylon staple in Table 1 show that increased twist does reduce the internal friction, but the results for continuous filament yarn suggest that this is mainly due to a reduction of fibre movement along the axis of the yarn under dynamic conditions. Insufficient data are at present available to correlate internal friction with elastic properties of the fibres. Under otherwise similar conditions a large elastic yield would be expected to produce a high internal friction.

Although in the theoretical treatment given the interfacial friction at the front edge appears from Eqn (5) to be less important than the work of deformation, this friction and the friction under the flat part of the slider is critically important in determining the value of p_e, i.e. the extent of any pressure build up that may occur. The interfacial friction is therefore important in so far as it affects the magnitude of the work of deformation. In the experiments described above, the surface of the front edge was comparatively rough, thus producing a rather larger μ and generally a very much larger excess edge friction than with a smoothly surfaced edge. For much larger values of μ or smaller values of β the interfacial edge friction may no longer be a negligible factor in Eqn (5).

It is clear that in many cases the internal friction may be of practical importance, for example, in increasing the total friction of a yarn running round a guide. The apparently greater internal friction of staple yarn as compared with continuous filament yarn and the correlation of this factor with liability to abrasion of the yarn may offer a direct approach to an important technological problem.

Acknowledgments

The author wishes to thank Mr. J. G. Hutton for experimental assistance, the referees for their helpful criticism, and the Director and Council of the British Jute Trade Research Association for permission to publish.

References

GREENWOOD, J. A., and TABOR, D., 1958, *Proc. Phys. Soc.*, **71**, 989.

HUFFINGTON, J. D., and STOUT, H. P., 1960, *Wear*, **3**, 26.

WALKER, A. C., and OLMSTEAD, P. S., 1945, *Text. Res. J.*, **15**, 201.

Narrow-beam echo-ranger for fishery and geological investigations

by M. J. TUCKER, B.Sc., F.Inst.P., and A. R. STUBBS, B.Sc., National Institute of Oceanography, Wormley, Surrey

MS. received 24th October 1960

Abstract

A 36 *kc/s underwater acoustic echo-ranger, fitted to R.R.S. 'Discovery II' for work on fish detection problems, has also proved to be a valuable tool for marine geological investigations. The axis of the acoustic beam is perpendicular to the fore and aft line of the ship and may be rotated from the horizontal to the downward vertical. Stabilization is provided against the roll of the ship. Beam width and side lobe patterns are adjustable with minimum widths, between* 3 *db points, of* 1·3° *in plan and* 12° *in elevation (beam axis horizontal). The recorder provides a range of* 800 *yards and scanning is achieved by steaming the ship past the features to be investigated.*

1. Introduction

For many years fishermen have used echo-sounders for locating fish and, more recently, horizontal echo-rangers to examine a larger volume of sea. There has been considerable discussion concerning the optimum parameters for such instruments, and the National Institute of Oceanography decided to build experimental equipment to investigate these.

Acoustic devices of many types have also been used to investigate the sea bed. One of particular interest was an echo-ranger which produced an acoustic picture of the sea floor (Chesterman 1958). Since this equipment was very similar to a horizontal fish finder, it was further decided to build the experimental equipment so that it could be used for both applications.

2. General principles

The equipment is fundamentally the same as an ordinary echo-sounder; a pulse of sound is sent out into the water and echoes from objects in its path are received and recorded on a chart or displayed on a cathode-ray tube. The time lapse between transmission of the pulse and the reception of the echo from a given target is a measure of its distance from the transmitter, that is, the range.

In the recorder, a stylus moves across a chart in a direction perpendicular to the chart's length, starting at one edge at the instant of transmission and marking the chart as the echoes are received. The chart moves slowly forward, so that each sweep is adjacent to the previous one and visual correlation of the echo patterns is possible.

Thus where the transducer beam is pointing sideways from a ship, the recorder will produce an acoustic map of the sea floor out to a range corresponding to the maximum travel of the recorder stylus. Also superimposed on the map will be objects (e.g. fish shoals) contained in the volume of water above this stretch of sea bed.

Echoes from near objects are much stronger than those from distant objects, and so the receiver gain must be arranged to be low immediately after transmission and to increase with time as the more distant echoes are received. The optimum law for this time-varied-gain, or 'T.V.G.', is not the same for all circumstances or for all purposes, and it is usually set empirically. It has been found advantageous to have a second parallel amplifier whose gain does not vary with time. This allows a greater flexibility in adjustment of the overall characteristics.

These parallel amplifiers are shown in the block diagram, Fig. 1. It will be seen also that the same transducer is used for both transmission and reception. Thus the cost is halved at the expense of a very slight reduction in performance (see Section 7).

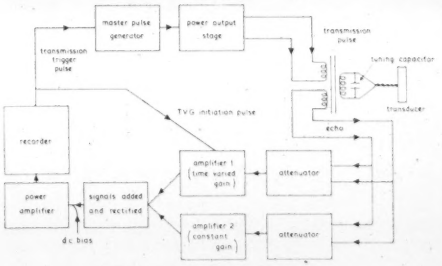


Fig. 1. Block schematic of system.

3. Choice of frequency

The range and acoustic beam widths are controlled, to a large extent, by the purpose for which the instrument is designed; these in turn largely govern the choice of frequency. There are, however, practical limitations and the final choice is generally a compromise governed by the following factors:

(i) For a given width of the sound beam the linear dimensions of the transducer are inversely proportional to the frequency, so that for low frequencies the size and cost rise steeply.

(ii) To get good resolution at long ranges, the beam must be narrow, but the length of the transducer is inversely proportional to beam width.

(iii) The higher the frequency the greater the sound absorption in water. Thus, the greater the range required, the lower the frequency that has to be used.

(iv) Echoes from the maximum range must be received before the next pulse is sent out. Thus the longer the range the less frequent the scans.

(v) It is an advantage to have a beam sufficiently wide to cover a small target for several scans, so as to distinguish it from noise or electrical interference.

(vi) The number of scans on a target is also dependent on the speed of the ship. If the speed is high the pulse repetition rate must also be high, and hence the range will be low. Thus, the slower the ship the better the coverage of targets. This in turn increases the cost of a given survey.

In the present case it was thought wise to aim at a beam width two to three times narrower than any at that time in use. This set a beam width of about 1°. It was considered that a range of about half a mile would be useful; this gives a pulse repetition rate of about one per second and it suggests a frequency in the range of 30–50 kc/s. It so happened that a large quantity of 36 kc/s magnetostriction elements were available and it was decided to make use of these and to design the system at this frequency.

4. The transducer

For both fishery and geological work, it was required that the vertical and horizontal beam patterns should be adjustable; the vertical pattern for the variation of side lobe intensity and the horizontal pattern for the variation of beam width.

Thus the transducer is divided into three rows of nine sections, each section consisting of thirty-two magnetostriction elements. The elements are made from Permalloy stampings and are polarized by a permanent magnet. The sections are mounted in juxtaposition horizontally, but are separated by 1·3 cm vertically, giving an active face of 156 cm × 21·5 cm. A photograph of the transducer assembly is shown in Fig. 2.

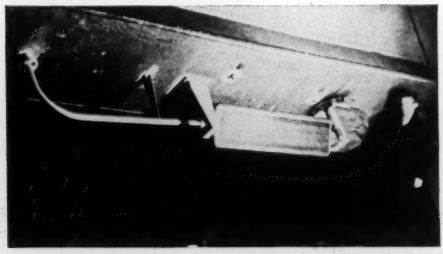


Fig. 2. Transducer assembly mounted on the hull of R.R.S. *Discovery II*

The sections are mounted in a bronze casting and protection is provided by a fibreglass window over the active face; the whole assembly is free flooding. The casting is mounted in trunion bearings held 14 in. from the hull on steel brackets bolted to pads welded to the hull 13 ft below the water line.

The beam axis can be rotated in a plane perpendicular to the ship's fore and aft line from inside the ship by a shaft passing through the hull and operating a worm and wheel. Electrical stops limit its movements to between +25 and −115° (zero being taken as horizontal). It is stabilized against roll by a system designed at Birmingham University (Tucker *et al.* 1959).

Each section is wound with a single length of polyvinyl chloride insulated wire, joined to two cores of a flexible fifty-four core cable, in a 'pickle-bottle' type junction box inside the after bearing. The cable leaves this axially and is protected by a hosepipe sleeving and supported by a steel tube until it enters the hull. The multicore cable terminates in a junction box where cross-connections are made to produce the desired beam patterns.

5. Acoustic properties

5.1 *Horizontal beam patterns.* When all nine columns of the transducer have the same sensitivity, a minimum beam angle of 1·3° between 3 dB points is obtained (see Fig. 3).

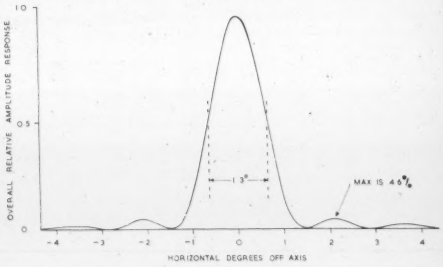


Fig. 3. Horizontal beam pattern (amplitude response).

Rectangular co-ordinates have been used to illustrate the relative amplitudes of the main beam and side lobes, the latter being less than 5% of the main beam for combined transmission and reception. As the side lobes are so small, horizontal tapering has not been used.

For most purposes such an arrangement has been employed, but when used in conjunction with the Birmingham University electronic scanning equipment (Tucker *et al.* 1959) one column was used solely for transmission to illuminate a sector 12° wide, and the other eight columns for reception with a beam width of about 1·6°.

5.2 *Vertical beam patterns.* Fig. 4 shows diagrammatically

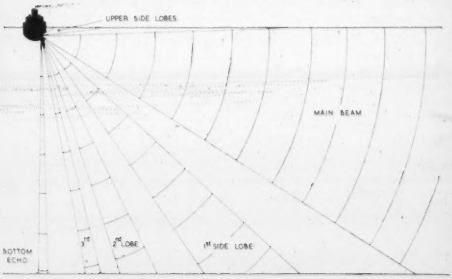


Fig. 4. Diagrammatic arrangement of vertical beam pattern.

the vertical beam pattern usually adopted. Below the main beam are three side lobes and a beam providing a bottom echo; the side lobes also occur above the main beam and strike the sea surface at short range. When all three rows of the transducer have the same sensitivity the main beam is about 12° deep, and this will increase to about 40° when only one row is used.

The vertical side lobes have been found to be extremely useful. Since there will be no bottom-returned echoes between them, features appearing in such positions on the record must be in mid-water and can be attributed to fish. These traces can then be used to identify fish shoals from bottom features, where these are superimposed. The lobes also provide a qualitative picture of the vertical relief of the sea bed, which helps in the interpretation of bottom features (see Fig. 14). Further, when the transducer is turned so that the first or second side lobes are pointing vertically downwards, the slope of the sea bed across the ship's track may be determined (Tucker, to be published).

For geological work it has been found advantageous to increase the amplitude of the first and third side lobes to obtain traces of comparable density. This is achieved by reducing the sensitivity of the centre row to two-thirds of that of the outer rows by a suitable series–parallel arrangement of the sections. The effect of 'tapering' can be seen in Fig. 5.

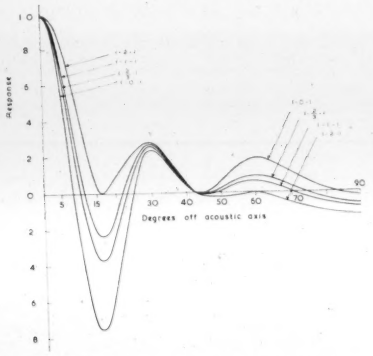


Fig. 5. Vertical beam pattern (amplitude response).

Two other sensitivity arrangements are also shown for comparison. By using a 1–2–1 arrangement, the first and third lobes are completely removed. This is useful on surveys for fish in the upper-water layers where bottom echoes are not required. (Main beam bottom echoes are removed by tilting the transducer up so that the lower edge strikes the sea floor beyond the maximum range of the recorder.)

In addition to the three secondary lobes already mentioned, there is also a small lobe between the second and third lobes. This is of such low amplitude that it does not produce any echoes and is ignored.

5.3 *Tilt angles.* To obtain the most efficient use of the chart width, the angle of depression of the beam below the horizontal has been found to be rather critical, and usually this is set at a value given by $1\frac{1}{2}°$ per ten fathoms of water depth. This puts the centre of the main beam on the sea floor at about the maximum range of the instrument. For very shallow water the angle has to be increased when sea surface echoes become severe owing to bad weather.

5.4 *Slant-range distortion.* The range of a target displayed on the recorder is, of course, the slant range, which is greater than the horizontal range, and this produces some distortion of the sea bed picture, especially at short ranges. The magnitude of the effect is easily obtained from Fig. 6.

5.5 *Effect of ship's speed.* The number of 'scans' on a given object will depend on its distance from the ship and

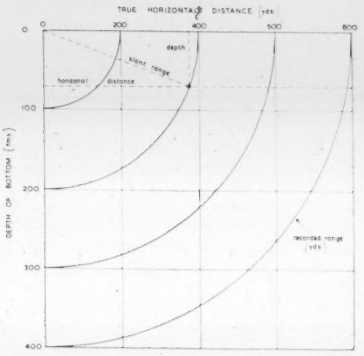


Fig. 6. Slant range as a function of depth and horizontal distance. $N = \frac{1}{25}\left(\frac{r}{K}\right)$.

on the ship's speed. For any horizontal beam width there is a theoretical limit to the ship's speed to obtain a reasonable coverage of the targets to be examined.

Fig. 7 shows the number of 'scans' on a point target for various ranges and ship's speeds (assuming a 1·3° beam width). It can be seen that for speeds of eight knots and under, all targets beyond about 200 yards range are given complete coverage.

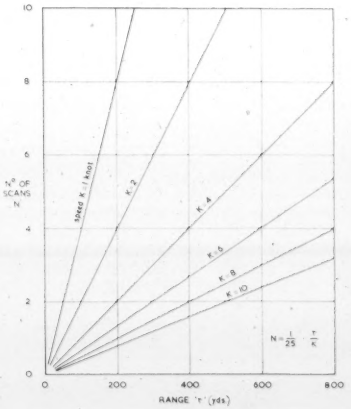


Fig. 7. The number of 'scans' on a point target as a function of speed and range.

6. Electronics

6.1 *Transmission.* The length of the pulse required was governed largely by the resolution of the recorder, and this was such that reduction of the pulse length below about 1 ms produced no detectable increase in recorded detail. Shorter pulses can be useful when a cathode-ray tube display is used, and longer pulses increase the signal-to-noise ratio. Thus the pulse was made to have its length adjustable between 0·3

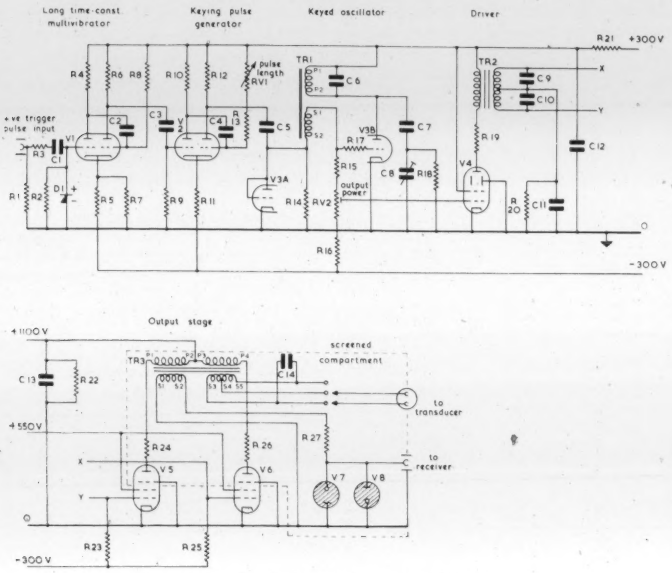


Fig. 8. Circuit diagram of transmitter.

R_1 = 220 kΩ
R_2 = 470 kΩ
R_3 = 22 kΩ
R_4 = 100 kΩ
R_5 = 150 kΩ
R_6 = 68 kΩ
R_7 = 22 kΩ
R_8 = 1·5 MΩ
R_9, R_{10} = 100 kΩ
R_{11} = 150 kΩ
R_{12} = 100 kΩ
R_{13} = 220 kΩ
R_{14} = 470 kΩ
R_{15} = 47 kΩ
R_{16} = 1 MΩ
R_{17} = 47 kΩ
R_{18} = 1 MΩ
R_{19} = 100 Ω
R_{20} = 330 kΩ
R_{21} = 22 kΩ
R_{22} = 1 MΩ
R_{23} = 680 kΩ
R_{24} = 22 Ω
R_{25} = 680 kΩ
R_{26} = 22 Ω
R_{27} = 100 kΩ
RV_1 = 500 kΩ
RV_2 = 100 kΩ
C_1 = 0·001 μF
C_2 = 0·5 μF
C_3 = 0·001 μF
C_4 = 0·01 μF
C_5 = 0·1 μF
C_6 = 680 pF
C_7 = 0·01 μF
C_8 = 96 pF
C_9, C_{10} = 0·003 μF
C_{11} = 1 μF
C_{12} = 8 μF
C_{13} = 16 μF
C_{14} = 0·01 μF
V_1, V_2 = 6J6
V_3 = 12AX7
V_4 = 6BW6 (CV4043)
V_5, V_6 = N339 (CV2231)
V_7, V_8 = LN1 (GEC)

and 3 ms, and to have a maximum repetition rate of two per second.

The circuit of the transmitter is shown in Fig. 8. The pulse is initiated when a pair of contacts in the recorder close, but, since they tend to 'bounce', the first closure is made to trigger a long time-constant monostable multivibrator, which is then insensitive to further closures for about 0·4 s. The pulse from this triggers a second monostable multivibrator, giving a square positive pulse of adjustable length which keys a 36 kc/s oscillator whose output passes to the power output stage.

The peak power output is 300 w, but, since the output stage does not need to deliver this continuously, the use of pulse modulator valves enables a compact design to be achieved. Those used in the present equipment are miniature valves, capable of passing a peak current of 2·5 A with a mean anode voltage of 6 kv. A pair of these in class C push–pull with an h.t. supply of 1·1 kv gives the required power. Similarly, if an adequate reservoir capacitor is used, only a low-power e.h.t. rectifier is required.

The output transformer of the transmitter is also the input transformer to the receiver. To prevent dangerously high voltage appearing in the cable and the receiver, neon lamps shunting the receiver fire during transmission. A series 100 kΩ resistor is used as a current limiter.

6.2 *The receiver.* The receiver is built in two parts for mechanical reasons, and the circuits of these are shown in Figs 9 and 10. The first part is a tuned 36 kc/s amplifier with a bandwidth of about 1 kc/s, the gain of which is low following transmission and rises with time until the next transmission. The output of the first stage is split into two channels, V11 has a gain which does not vary with time, V10 has a gain controlled by the voltage on C32. The outputs of these two valves are recombined in a common anode load, further amplified by V13 and fed to the second part of the receiver.

The time-controlled gain is initiated by the closure of the transmission contacts on the recorder. At this instant the high-speed relay P releases and the contact P1 causes a negative charge to be put on C32, which reduces the gain of the variable-mu pentode V10. The gain recovers as the charge leaks away through RV4.

The variation of gain with time is thus governed by four controls: the gain controls of the two channels and the settings of RV3 and RV4. In practice, RV3 and RV4 are set so that, roughly speaking, switch SW2 controls the density of the marking on the close-range portions of the record, and switch SW1 controls the density of marking of the long-range parts of the record. A graph of the variation of gain with time at typical settings is shown in Fig. 11.

It is very convenient to have reference marks on the record, either as time marks or for correlating features with other recorders. Relay MR performs this function and may be operated by a push-button on the panel or remotely by

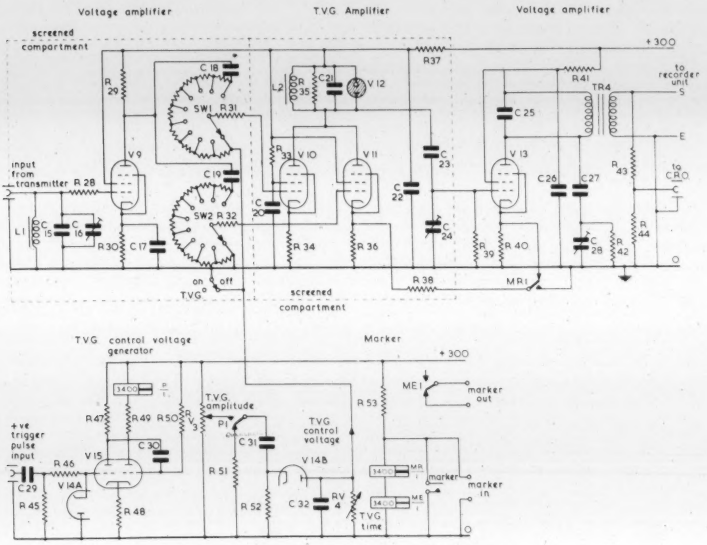


Fig. 9. Circuit diagram of receiver amplifier.

$R_{28} = 47\ k\Omega$
$R_{29} = 22\ k\Omega$
$R_{30} = 330\ k\Omega$
$R_{31}, R_{32} = 1\ k\Omega$
$R_{33} = 47\ k\Omega$
$R_{34} = 68\ k\Omega$
$R_{35} = 220\ k\Omega$
$R_{36} = 330\ k\Omega$
$R_{37} = 2 \cdot 2\ k\Omega$
$R_{38} = 33\ k\Omega$
$R_{39} = 330\ k\Omega$

$R_{40} = 470\ \Omega$
$R_{41} = 10\ k\Omega$
$R_{42} = 330\ k\Omega$
$R_{43} = 330\ k\Omega$
$R_{44} = 10\ k\Omega$
$R_{45} = 100\ k\Omega$
$R_{46} = 47\ k\Omega$
$R_{47} = 100\ k\Omega$
$R_{48} = 2 \cdot 2\ k\Omega$
$R_{49} = 22\ k\Omega$
$R_{50} = 1\ M\Omega$

$R_{51} = 6 \cdot 8\ k\Omega$
$R_{52} = 150\ k\Omega$
$R_{53} = 33\ k\Omega$

$RV_3, RV_4 = 500\ k\Omega$

$C_{15} = 470\ pF$
$C_{16} = 96\ pF$
$C_{17} = 0 \cdot 5\ \mu F$
$C_{18}, C_{19} = 0 \cdot 001\ \mu F$
$C_{20} = 0 \cdot 1\ \mu F$

$C_{21} = 680\ pF$
$C_{22} = 2\ \mu F$
$C_{23} = 0 \cdot 001\ \mu F$
$C_{24} = 96\ pF$
$C_{25} = 470\ pF$
$C_{26} = 0 \cdot 1\ \mu F$
$C_{27} = 0 \cdot 001\ \mu F$
$C_{28} = 96\ pF$
$C_{29} = 0 \cdot 01\ \mu F$
$C_{30} = 0 \cdot 1\ \mu F$
$C_{31}, C_{32} = 1\ \mu F$

$V_9 = EF91$
$V_{10} = 6BA6$
$V_{11} = EF91$
$V_{12} = LN1\ (GEC)$
$V_{13} = EF91$
$V_{14} = 6AL5$
$V_{15} = 6J6$

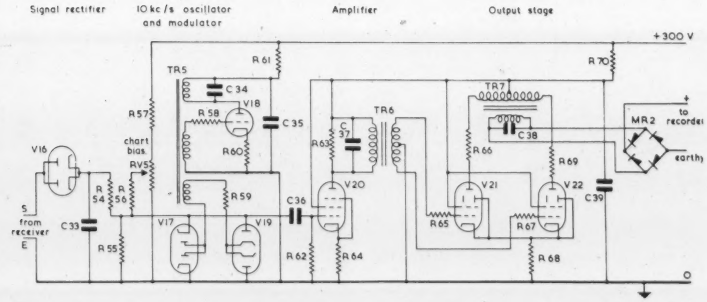


Fig. 10. Circuit diagram of recorder output unit.

$R_{54} = 150\ k\Omega$
$R_{55} = 6 \cdot 8\ k\Omega$
$R_{56}, R_{57} = 1\ M\Omega$
$R_{58} = 10\ k\Omega$
$R_{59} = 4 \cdot 7\ k\Omega$
$R_{60} = 1\ k\Omega$

$R_{61} = 33\ k\Omega$
$R_{62} = 1\ M\Omega$
$R_{63} = 68\ k\Omega$
$R_{64} = 680\ \Omega$
$R_{65} = 10\ k\Omega$
$R_{66} = 100\ \Omega$

$R_{67} = 10\ k\Omega$
$R_{68} = 330\ \Omega$
$R_{69} = 100\ \Omega$
$R_{70} = 470\ \Omega$

$RV_5 = 100\ k\Omega$

$C_{33} = 0 \cdot 005\ \mu F$
$C_{34} = 0 \cdot 03\ \mu F$
$C_{35} = 2\ \mu F$
$C_{36} = 220\ pF$
$C_{37} = 0 \cdot 0022\ \mu F$
$C_{38} = 0 \cdot 2\ \mu F$
$C_{39} = 2\ \mu F$

$V_{16}, V_{17} = 6AL5$
$V_{18} = 6C4$
$V_{19} = 6AL5$
$V_{20} = EF91$
$V_{21}, V_{22} = 6BW6\ (CV4043)$

short-circuiting the remote marker leads. Closure of the MR1 contact connects positive feedback round the circuit, including V11 and V13, which then oscillate to produce a black line on the record. Relay ME enables simultaneous marks to be put on any other equipment.

The second part of the receiver circuit is arranged to add a

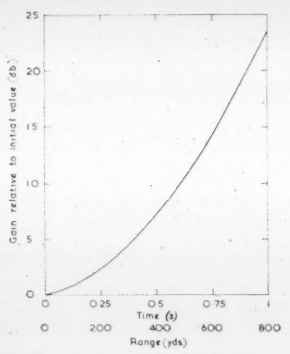


Fig. 11. Graph showing a typical variation of amplifier gain with time following transmission.

bias to the signal and then amplify it to mark the recorder chart. The 36 kc/s signal is rectified, the d.c. bias added through RV5, remodulated at 10 kc/s, amplified, rectified and fed to the recorder. The purpose of the biasing is described in the next section.

6.3 *The recorder*. This is a slightly modified D–611–A Mufax picture receiver (by Muirhead and Co. Ltd.) with an 11 in. wide chart. The stylus is a platinum alloy wire wound as a one-turn helix on a drum. The helix is driven by a synchronous motor controlled by a tuning fork, precise control of speed being necessary to avoid distortion of the picture. The transmission pulse is initiated by a micro-switch operated by a cam on the end of the helix-drum driving shaft.

The chart passes between the helix and a knife-edge parallel to the axis of the drum, so that the point of contact moves along the knife-edge as the drum rotates. For the present requirement the chart is driven so that there are fifty scans per inch, at which speed individual scans can just be resolved but a coherent picture is produced. The resolving power of the recorder is between 0·01 and 0·02 in., which at one scan per second is equivalent to between 1 and 2 ms, or 5–10 ft in range.

The biasing carried out in the recorder output unit is required because of the characteristics of the wet electrolytic recorder paper. Passing a current through the stylus electrolyses the solution impregnating the paper and, at the same time, releases some iron from the stylus blade. A chemical reaction takes place, which then produces a black mark on the paper.

About 6 v are required before any noticeable colour is produced, and about 20 v produces full blackness. Thus in photographic terms, the paper is 'hard' and detail is lost. To 'soften' it, a bias is required and it has been found that the method used to introduce it, though apparently complicated, is the most satisfactory.

7. Sensitivity, noise and calibration

The sensitivity of the system is limited by noise. Considerable difficulty was experienced due to interference from the ship's generators. The transducer is closely coupled to the sea, both capacitatively through the windings, which are surrounded by sea water, and through a certain amount of resistive leakage. The best results were obtained by leaving the transducer circuit floating with respect to earth, it being isolated from the electronics by the input transformer, and connecting the electronics to an earthing bolt on the hull by a short thick wire.

Owing to the high cost of the transducer, the same one is used for both transmission and reception; this has a negligible effect on the performance, lowering the signal-to-noise ratio by about 1 dB compared with the use of two transducers.

With the ship under way, the limiting noise was then due to air bubbles passing the transducer face. This depended very much on ship's speed and weather conditions. In good conditions a signal of 1 μv at the transducer terminals could be detected above noise (transducer impedance $9 + 25j\Omega$). This corresponds to the detection of a perfectly reflecting sphere with a radius of 10 cm at a range of 500 m.

Under bad conditions, noise could be 15 dB worse than this and, in addition, the aeration in the water increased the attenuation enormously. It is usually impossible to obtain satisfactory results in winds of over 25 knots.

The calibration of the transducer was performed by measuring the echo strengths from targets of known acoustic cross section, and by measuring the electrical impedance at various frequencies both in air and in water. The two methods agreed within about 25%, which is satisfactory considering the uncertainties involved.

8. Interpretation of records

Figures 12–15 are typical records, each 800 yards wide and about two to three miles long; the transmission mark corresponding to zero range is on the left-hand side. Interpretation of the records should be made in conjunction with Fig. 4.

The first series of echoes are from the sea surface, produced by the upper side lobes, and are displayed quite close to the transmission mark. Ignoring mid-water echoes, the next echo to arrive is that from the sea bed directly under the ship. Though the transducer is very insensitive in this direction, the reflecting power of the sea bed at normal incidence is high enough to produce a reasonably strong echo, which appears as a hard line on the record.

This is followed quickly in time by echoes from the third side lobe, then at wider intervals by echoes from the second and first side lobes. All these echoes are quite strong, since they come from shorter ranges and the beams have a larger angle of incidence to the sea bed. Finally, echoes are received from the sea bed illuminated by the main beam. These will be displayed across the record at ranges depending on the depth of water and the tilt of the transducer. The records are, in effect, acoustic charts of the sea bed with mid-water echoes from fish shoals, etc., superimposed.

These 'charts' are not true to scale, however, distortion being produced by a variety of causes. The chart width is 800 yards, but this is the slant distance to any target at that range. The distortion is small in shallow water but increases with depth. This is, of course, true for all points on the record, the distortion being greater as the range decreases (see Section 5.4). The chart is also compressed along its length

compared with its width, typically by about 3 to 1. This ratio will depend on the ship's speed and the chart speed.

A further correction has to be applied when the ship's heading does not agree with the course made good over the ground, owing to strong wind or water movements. Under these conditions the beam will not be perpendicular to the ship's track. In general, this is not a serious effect, but is of importance when detailed surveys are carried out on exposed coasts or in areas with strong tidal streams. The effect produced by the roll motion of the ship when the transducer is

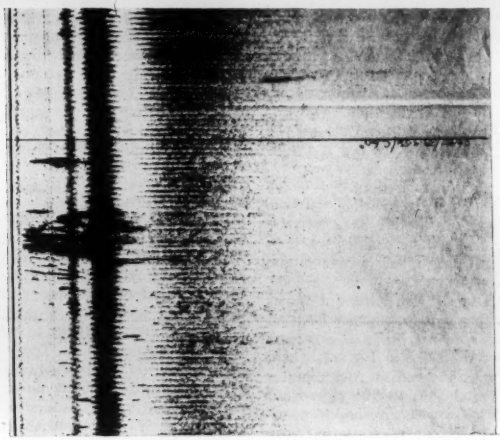


Fig. 12. Record showing fish concentration over a flat sandy sea floor. Transducer unstabilized and no time-varied-gain.

Fig. 13. Record showing a large stratified slate formation.

Fig. 14. Record showing a large area of stratified rock layers.

unstabilized is indicated by the saw-tooth patterns of Fig. 12. This is removed almost completely by the roll stabilization system, as shown in the subsequent figures. Distortion also occurs due to the ship pitching, yawing and heaving, but these are second-order effects and are usually ignored.

Figure 12 shows a record having fish shoals over a flat sandy sea bed. The main concentration is obviously in mid-water, since the nearest echoes are received before the bottom echo. Small strong echoes towards the bottom of the record, appearing between the side lobe echoes, are also fish in mid-water, detected by the main beam. These cannot be bottom echoes since the side lobes are ineffective at these points on the sea bed. At shorter ranges fish in the main beam give strong echoes relative to the reverberation from the sea bed in the side lobes; at greater distances the sea bed is also in the main beam and fish shoals do not stand out so well.

The picturesque echo patterns in the centre of Fig. 13 are produced by a large stratified slate formation, which has been fractured and tilted. The dark lines are produced by slate edges facing the transducer and the white patches and lines are shadow zones behind these edges. The bottom profile shows the slate to be protruding from a fairly flat sea floor, which is probably composed of sand. The pattern of dashes on the right-hand side of the record is interference from an echo-sounder.

Layers of stratified rock are exhibited in Fig. 14, where there are also distinct fault lines. The beds of rock dip at a few degrees below the horizontal towards the right-hand side of the record. The large white areas are shadow zones behind the rock layers. Some of the echoes, appearing close to the bottom echo and displayed across the side lobes, are probably from fish, but the majority are from a confused sea surface due to bad weather.

A complex sand-wave area is depicted in Fig. 15. The gentle slopes of the widely spaced sand waves are toward the transducer, the crests are about 40–100 yards apart. Superimposed on these are smaller sand waves of about 5–10 yards separation. The vertical height of the bigger sand waves is not very large, probably not greater than 10 ft; this is indicated by the bottom profile. The darkened portion between the transmission and the bottom profile is caused by echoes from a confused sea surface due to bad weather; interference from an echo-sounder is again evident at the farthest ranges.

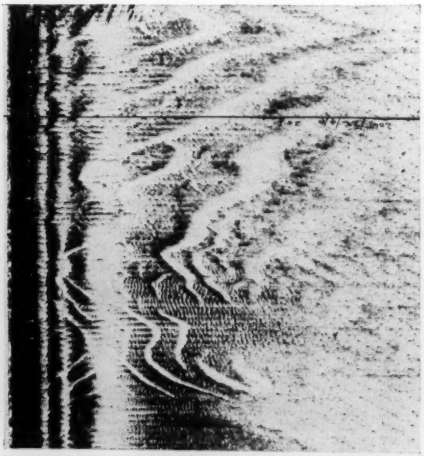


Fig. 15. Record showing a complex sand-wave area.

Interesting records have also been obtained from a variety of objects; these have included waves, ships' wakes, wrecks, cables, piers and porpoises. The latter, in addition to providing an echo in the usual manner, also emit pulses of high-frequency sound, which, being detected by the transducer, are displayed randomly as noise across the record.

9. Applications

The most extensive use of the instrument up to the present time has been for studying the geology of the sea bed, both for purely geological purposes (Stride 1959, 1961) and for studying the relationship of fish shoals with their habitat (Stubbs and Lawrie, to be published). It has been used in conjunction with other equipment by Harden Jones (discussion on Tucker 1960) to study the movement of fish relative to tidal streams. Trials have been made to evaluate its potentialities for charting fish distribution.

When turned vertically downwards it forms a powerful tool for studying the sound-scattering layers in the sea, since owing to its high resolution, it gives a great deal more information than a wide-beam sounder. For example, some layers which appear to be continuous on an ordinary sounder are resolved into individuals, or sometimes into small groups of scatterers. Preliminary attempts to correlate the records obtained with the net hauls have had limited success, however.

In some circumstances, the high resolution obtained when used as a vertical sounder can be useful for studying the topography of the sea bed, but the range is too short for most deep-sea work.

The limiting factors for all applications are water noise and aeration produced at high speed or during bad weather. A towed transducer would tend to overcome these, but the problem of roll stability in this case would call for careful design. The use of side lobes has been shown to be advantageous and future equipment may benefit by increasing their number.

Acknowledgments

The authors wish to acknowledge the assistance given by Mr. F. E. Pierce of the Institute design section, together with numerous colleagues who have given help in the design and production of the instrument.

References

Chesterman, W. D., Clynick, P. R., and Stride, A. H., 1958, *Acustica*, **8**, 285.

Stride, A. H., 1959, *J. Mar. Biol. Ass. U.K.*, **38**, 313.

Tucker, D. G., 1960, *J. Brit. Instn Radio Engrs*, **20**, 299.

Tucker, D. G., Welsby, V. G., Kay, L., Tucker, M. J., Stubbs, A. R., and Henderson, J. G., 1959, *J. Brit. Instn Radio Engrs*, **19**, 681.

Spectral emissivity of 99·7% aluminium between 200 and 540° C

by P. M. REYNOLDS, Ph.D., British Non-Ferrous Metals Research Association, London, N.W.1

MS. received 7th November 1960

Abstract

Results are given for surfaces of roughness 3 μ *in.*, 115 μ *in. and anodized aluminium within the wavelength range* 1 *to* 14 μ. *The spectral emissivity varied strongly with surface roughness and oxidation, but only slightly with temperature. Typical values at wavelengths* 2 *and* 8 μ *respectively were* 0·09 *and* 0·04 *for polished*, 0·30 *and* 0·24 *for roughened*, 0·30 *and* 0·20 *for anodized aluminium. Natural oxidation slightly increased the spectral emissivity at all wavelengths below* 9 μ *and had a marked effect between* 10 *and* 12 μ. *An anodic oxide film increased the spectral emissivity beyond* 10 μ *to greater than* 0·70. *The significance of these variations to radiation pyrometry is briefly discussed.*

1. Introduction

An accurate method of measuring the temperature of aluminium and its alloys during hot working (400–550° C) would be of considerable industrial importance. Thermocouple contact pyrometers have a low speed of response (8 to 20 s), cannot generally be applied to rapidly moving surfaces and some designs damage the surface finish. Radiation pyrometers have the advantages of rapid response (5 to 10^{-3} s) and of not requiring contact with the specimen, but the disadvantages that their output depends upon both the temperature and emissivity of the specimen. Corrections for the emissivity may be applied by calibration against a specimen of similar emissivity, by using methods of increasing the effective emissivity of the specimen (Harmer and Watts 1955, Land and Barber 1954), or by assuming a known variation of spectral emissivity with wavelength and using two- (Gibson 1951) or three-colour (Brendon and Newkirk 1959) pyrometers. With the exception of two- and three-colour pyrometers, instruments employing these methods of emissivity correction and suitable for this temperature range are commercially available. With metals of low and variable emissivity, such as aluminium, it is essential to know the variation of spectral emissivity with wavelength before the most suitable method of emissivity correction can be selected.

2. Experimental method

The spectral emissivity was determined directly at known wavelengths by comparing the radiation emitted from an aluminium specimen with that from a black-body source at the same temperature. The specimen A, Fig. 1, was a thin-walled cylinder of length 5·5 cm, diameter 1·2 cm, and wall hickness 0·003 cm, which was directly heated by a current of 80–150 A. Radiation emerging through the rectangular slit S, 0·1 cm × 1·2 cm, approximated to black body radiation at the same temperature as the external surface of the tube. At C the specimen was clamped to cylindrical steel terminals T, each heated by a subsidiary winding H. The temperature distribution within the black-body cavity could be explored with a thermocouple Th, and the subsidiary heaters adjusted to compensate for end-cooling of the specimen. The temperature variation over the central 3 cm of the specimen did not exceed ±1 deg C.

Figure 2 shows the general arrangement of the apparatus. The specimen was exposed to the atmosphere and surrounded

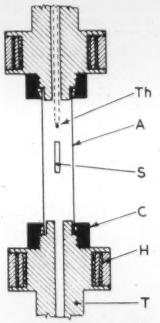


Fig. 1. Specimen and mount.

by a water-cooled jacket J, which contained an aperture through which the specimen could be examined. The internal surface of this water jacket was blackened. Radiation from the specimen was condensed upon the entrance slit of a monochromator (Hilger type D246) containing a rock-salt prism. The incident radiation was chopped at 10 c/s by rotating blades B and, after dispersion, focused upon a Golay detector G (Unicam). A beam-switching mirror M enabled the radiation from either the black-body slits, the aluminium surface adjacent to the slit or a zero reference black-body Z in the same optical plane as the specimen, to be selected. The zero reference black-body was water-cooled and contained a blackened cavity 2·0 cm × 1·2 cm with slit 0·1 cm × 1·2 cm to match the cavity in the specimen. Its function will be described later.

3. Principle

Let T_a, T_d, T_w and T_z be the absolute temperatures of the specimen, detector, water jacket and zero reference black-body respectively, and E_a, E_d, E_w and E_z be the energy flux radiated by a black-body at these temperatures per unit area at wavelength λ for a wavelength interval $d\lambda$. The

spectral emissivity of the aluminium at wavelength λ and temperature T_a is ϵ_λ. The detector, water jacket and zero reference black-body are assumed to have a spectral emissivity equal to unity.

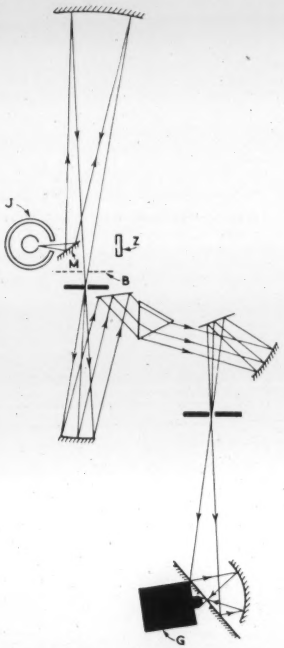


Fig. 2. Optical system.

Neglecting multiple reflection, the energy flux W_a sensed by the detector when sighted upon the aluminium is given by

$$W_a = \{\epsilon_\lambda(E_a - E_d) + (1 - \epsilon_\lambda)(E_w - E_d) + X\}\phi \quad (1)$$

where ϕ is a geometrical factor. The first term represents the direct energy exchange between aluminium and detector. The second term allows for radiation from the surrounding water jacket which reaches the detector after reflection from the aluminium, and X is the radiation from the chopper blades. For measurements at high temperatures with specimens of high emissivity the last two terms are insignificant compared with the first and $E_a \gg E_d$. Eqn (1) then reduces to

$$W'_a = \epsilon_\lambda E_a \phi. \quad (2)$$

Sighting upon the heated black-body source, the energy flux collected by the detector becomes

$$W_b = \{(E_a - E_d) + X\}\phi. \quad (3)$$

At high temperature this reduces to

$$W'_b = E_a \phi. \quad (4)$$

Thus at high temperature or with specimens of high emissivity, ϵ_λ can be determined simply from the ratio W'_a/W'_b.

The spectral emissivity of polished aluminium at wavelengths greater than 6 μ lies between 0·03 and 0·06. At low temperatures the terms $(1 - \epsilon_\lambda)(E_w - E_a)$ and X in Eqn (1) become significant. For example, when $T_a = 200°$ C, $\lambda = 8\ \mu$ and $\epsilon_\lambda \simeq 0{\cdot}03$, a 1 deg C change in T_w, the temperature of the water jacket, altered W_a by 4%; W_b was constant. The value of ϵ_λ determined from the ratio W_a/W_b then varied with the temperature of the water jacket, detector and chopper blades. To eliminate this unwanted radiation it was necessary to measure a difference signal. A zero reference black-body was used. The energy flux sensed by the detector from this source is given by

$$W_z = \{(E_z - E_d) + X\}\phi. \quad (5)$$

When $T_w = T_z$, the signals $W_a - W_z$ and $W_b - W_z$ reduce to

$$W_a - W_z = \epsilon_\lambda(E_a - E_z)\phi \quad (6)$$

and

$$W_b - W_z = (E_a - E_z)\phi \quad (7)$$

and ϵ_λ can be determined from their ratio.

4. Accuracy and reproducibility

To ensure that the reflected radiation arose entirely from the water jacket, and not from room temperature radiation specularly reflected through the aperture in the water jacket, it was necessary to collect radiation at about 15° from the normal to the specimen surface. At 200° C, when $\lambda = 8\ \mu$, a 1 deg C change in T_w then altered ϵ_λ by 1%. Recently it has been found that the inner surface of the water jacket was not perfectly black. When this was remedied a 10 deg C change in T_w altered ϵ_λ by 2%. The actual value of T_w (and T_z) during these experiments was $(19 \pm 2)°$ C.

With roughened or anodized aluminium, when $\epsilon_\lambda > 0{\cdot}20$, the errors due to stray signals were negligible.

At any given temperature, values of the spectral emissivity for a given area of each specimen were reproducible within $\pm 2\%$ near the peak of the emission curve and within $\pm 10\%$ at the short and long wavelength limits. In addition the spectral emissivity varied over the surface of each specimen due to small variations in surface finish. At 3·5 μ this variation was about $\pm 2\%$ for all specimens. The calculated efficiency of the heated black-body was 97·5% (De Vos 1954).

The estimated accuracy over the wavelength range 2 to 10 μ was $\pm$ 20% for polished aluminium and $\pm 10\%$ for roughened or oxidized aluminium. Outside this range the accuracy was lower.

5. Preparation of specimens

The specimens were formed from extruded commercial purity aluminium containing the following impurities: Fe, 0·11%; Si, 0·11%; Cu, 0·01%; Mg, 0·01%; Mn, Ni and Zn, $< 0{\cdot}01\%$. Three different surface finishes were studied (i) high polished, (ii) roughened, and (iii) anodized. The natural surface of the extruded tube was smooth and a high polish was obtained using Carnu (by S. C. Johnson Ltd.) polish on Selvyt cloth. The surface roughness was 3 μ in. (centre line average (c.l.a.), sampling length 0·01 in.). The rough aluminium tube had a surface roughness 115 μ in. c.l.a. (sampling length 0·03 in.) and was prepared by mounting the tube on a mandrel and knurling its surface with grade 180 silicon carbide paper attached to a rotating wheel. This produced an imprint of the silicon carbide paper on the aluminium. After washing there was no

evidence of embedded silicon carbide when the specimen was examined with a low-power (×30) microscope. To prepare the anodized surface, a specimen was anodized for 30 min at 1 A/dm^2 in a bath of 4 N analar sulphuric acid which was maintained at 20° C and vigorously agitated, and was finally sealed for 30 min in boiling distilled water. The thickness of the anodic film was determined by a micro-section and found to be 0·0001 in.

6. Experimental results

Figure 3 shows the variation of spectral emissivity with wavelength for the polished tube as determined using Eqns (6) and (7). The tube was maintained at 194° C for 15 h, 326° C

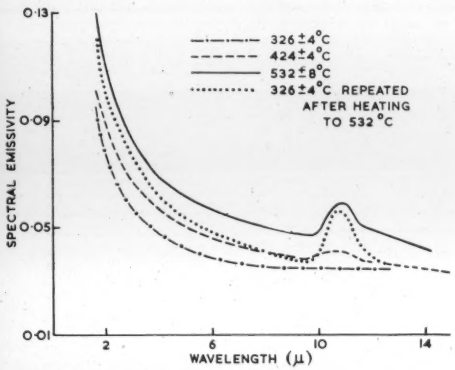


Fig. 3. Spectral emissivity of polished aluminium plotted against wavelength.

for 19 h, 424° C for 20 h and 532° C for 15 h. Finally, the tube temperature was reduced to 326° C and the spectral emissivity redetermined. As the specimen temperature increased the thickness of the natural oxide film also increased, and a peak appeared between 10·5 and 11·0 μ. This was more pronounced after the specimen had been heated to a higher temperature, but oxidation increased the spectral emissivity at all wavelengths. The increase in spectral emissivity with temperature was primarily due to the growth of an oxide film.

The spectral emissivity of the rough tube is shown in Fig. 4. This tube was heated for 25 h at 189° C, 22 h at 325° C, 27 h at 442° C, and 17 h at 514° C. The temperature was then reduced to 192° C and further results obtained. Roughening the specimen strongly increased the spectral emissivity at all wavelengths and emphasized the peak value between 10·5 and 11·0 μ. Subsidiary peaks appeared around 8·5 and 6·0 μ as oxidation proceeded. After prolonged heating at a higher temperature the spectral emissivity at 192° C was considerably increased at wavelengths beyond 10 μ.

Comparing these results with Fig. 5, which shows the spectral emissivity of the anodized tube, it is clear that the maxima at 8·5, 10·5 and 11·0 μ must be due to an oxide film. Additional maxima occur at 2·7 and 4·7 μ which were not observed on naturally oxidized specimens. The sequence of heating for the anodized tube was 15 h at 183° C and 8 h at 302° C.

7. Comparison with results of previous work

The optical properties of anodized aluminium oxide films stripped from their metallic base have been studied by Harris (1955), who found strong absorption beyond 11 μ for films as thin as 250 Å. Hase (1932) determined the spectral emissivity of 99·6% aluminium at 300, 400, 500 and 600° C over the wavelength range 1 to 9 μ and the effect of surface finish and oxidation. His results showed that the spectral emissivity of polished aluminium varied irregularly with wavelength and temperature between 0·07 and 0·11, except below 2·5 μ when the emissivity increased rapidly with decreasing wavelength. Roughening the aluminium

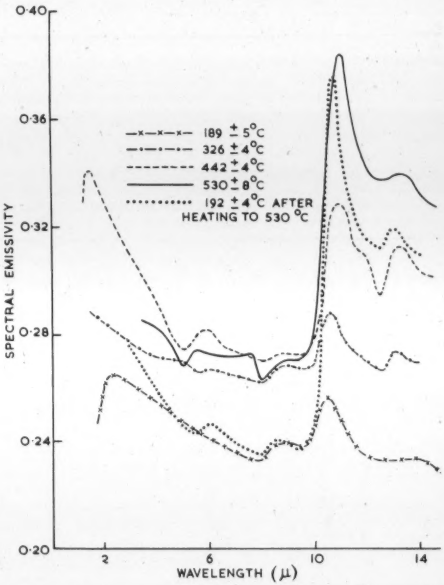


Fig. 4. Spectral emissivity of roughened aluminium plotted against wavelength.

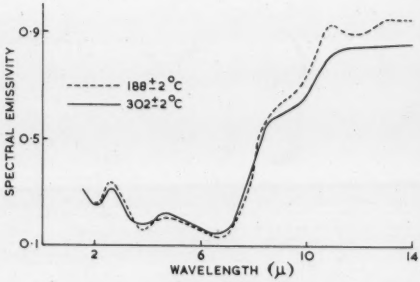


Fig. 5. Spectral emissivity of anodized aluminium plotted against wavelength.

surface strongly increased the spectral emissivity; values as high as 0·7 were obtained at 2 μ decreasing to 0·42 at 8 μ. For heavily oxidized surfaces peak values occurred at approximately 2, 3, 6 and 9 μ. More recently, Weber (1959) investigated the spectral emissivity of aluminium alloys type 24ST (nominal composition Cu 4·5%, Mn 0·6%, and Mg 1·5%) and 76ST (Cu 1·6%, Mn 0·2%, Mg 2·5%, Zn 5·6% and Cr 0·3%) between 30 and 110° C for wavelengths 7 to 15 μ. The spectral emissivity did not exceed 0·05 for smooth samples selected from stock. These results agree qualitatively with the present work.

8. Conclusion

The spectral emissivity of commercial purity aluminium at temperatures between 200 and 540° C and wavelengths 2 to 15 μ varies considerably with wavelength, surface roughness and state of oxidation. Roughening the surface increases the spectral emissivity at all wavelengths, whereas oxidation tends to increase the spectral emissivity at selected wavelengths. The latter effect is particularly marked at wavelengths beyond 10 μ.

9. Application to radiation pyrometry

These results indicate that, even if calibrated against an aluminium surface, a radiation pyrometer may be subject to large errors due to variations in emissivity. These errors will depend upon the range of wavelengths accepted by the detector. For example, a total radiation pyrometer sensitive beyond 9 μ will be more liable to errors due to variations in surface roughness and oxidation than a pyrometer only sensitive below 9 μ. At temperatures below about 300° C a pyrometer sensitive beyond 6 μ may also respond to variations in ambient radiation reflected from the aluminium surface whenever the spectral emissivity is low. Since the spectral emissivity of aluminium is not independent of the wavelength the condition for the use of a two-colour pyrometer cannot be satisfied. For a three-colour pyrometer the spectral emissivity must vary linearly with wavelength; again this requirement is not fulfilled by aluminium. The alternative methods, which increase the effective emissivity of the specimen and decrease the relative variations in effective emissivity with surface condition, either by using methods of multiple reflection (Land and Barber 1954) or by employing detectors such as the lead sulphide cell (Harmer and Watts 1955) appear more promising and will be discussed elsewhere.

References

Brendon, B. B., and Newkirk, H. W., 1959, *Unclassified report HW*-60678, *UC*–37, *Instruments*, (Washington: Hanford Atomic Products Operations).

De Vos, J. C., 1954, *Physica*, **20**, 669.

Gibson, A. F., 1951, *J. Sci. Instrum.*, **28**, 153.

Harmer, J. D., and Watts, B. N., 1955, *J. Sci. Instrum.*, **32**, 167.

Harris, L., 1955, *J. Opt. Soc. Amer.*, **45**, 27.

Hase, R., 1932, *Z. Tech. Phys.*, **13**, 145.

Land, T., and Barber, R., 1954, *J. Soc. Glass Technol.*, **38**, 45.

Weber, D., 1959, *J. Opt. Soc. Amer.*, **49**, 815.

Heating of metallic foils in an electron microscope

by B. GALE, B.Sc., Ph.D., and K. F. HALE, B.Sc., A.Inst.P., National Physical Laboratory, Teddington, Middlesex

MS. received 16*th September* 1960, *in final form* 8*th November* 1960

Abstract

The temperature of a thin metal foil illuminated by the high intensity electron beam of the modern microscope may at the centre of illumination reach a temperature above that of the melting point of iron. An attempt has been made to measure and calculate the temperature distribution for any given metal under various illumination conditions and to determine also the stress fields produced in the foil by such temperature distribution.

DIRECT observation of dislocations and other imperfections in metals has, in recent years, been mainly carried out using electron microscopy. In this technique foils are prepared having a thickness between 1000 and 4000 Å (Carrington *et al.* 1960) and are examined at magnifications of the order of $\times 10^5$, using electrons accelerated by voltages up to 100 kv. The specimens are thus irradiated by an intense electron beam which, as will be shown, can raise the temperature considerably. This, together with the ensuing thermal stresses can produce large effects (Whelan 1957) superimposed on the relatively precise thermal and mechanical treatments given to the specimens prior to examination.

The heat generated in the specimen is proportional to the incident electron current distribution which was determined experimentally, on a relative scale, for a Siemens Elmiskop I microscope. With this instrument the size of the irradiated area, and consequently the heat input, can be varied using the first and second condenser lenses. The microscope was operated with an accelerating voltage of 100 kv, with a beam current of 5 μA at a magnification of $\times 4600$, the filament voltage being kept constant at 2·7 v and a 30 μ condenser aperture being inserted to reduce the illumination to a manageable level. A characteristic curve for Ilford special lantern contrasty plates was determined for this voltage by photometering exposures of different duration, precautions being taken to allow for changes in the beam intensity during the experiment. Constant time exposures were then made of the illumination for the whole available range of the first condenser current. In each case the second condenser was focused to give minimum irradiated area. It was necessary to adjust the condenser traverse and gun centring controls to give maximum brightness and, as before, to allow for drift in the beam intensity. Using the characteristic curve just determined it was found that the intensity distributions were symmetrical about the beam axis and could be represented by the Gaussian relation $I(r) = I_0 \exp\{-(r/a)^2\}$, where r is the distance from the beam axis, a is a measure of the extent of the irradiated area, determined by the currents in the condenser lenses, and I_0 is the peak intensity. It was found that I_0 was constant within the experimental error of about $\pm 5\%$, as has been found by the Siemens Laboratories (Wolff, private communication).

Direct measurement of the temperature rise produced in the specimen by the absorption of the electron beam does not seem possible because of the small size of the irradiated volume (between 10^{-12} and 10^{-9} cm^3). Instead, the temperature was calculated as a function of the parameter a, with a constant of proportionality determined by observing the values of a at which known physical changes occur in various materials. In the microscope the specimen foils are laid on to a copper grid, but they are not closely in contact with it in any predetermined way. As determined experimentally the electron irradiation of the foil is circularily symmetrical, but it is not necessarily centrally situated with respect to any one of the square grid frames. Thus, in view of the considerable complexity of the situation, it was thought advisable to replace it by a simplified model which is reasonably close to it.

In this model it is assumed that the foil is bounded by a circular conductor of infinite conductivity held at a fixed temperature T_0, and that the irradiation intensity is symmetrical about the boundary centre. The rate of generation of heat per unit volume $H(r)$ is assumed to be independent of the depth of penetration of the electron beam, and is thus solely a function of the distance r from the centre of the beam. The temperature T in the foil is thus only a function of r, and is determined by the solution of the two-dimensional heat conduction equation, which for this problem is of the form:

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dT}{dr}\right) + \frac{1}{k}H(r) = \frac{2\sigma}{kd}(T^4 - T_0^4) \qquad (1)$$

where k is the thermal conductivity, σ is Stefan's constant and d is the thickness of the foil. Von Borries and Glaser (1943) have calculated solutions of this equation for the special case when $H(r)$ is constant and the conductivity is small, whereas Whelan *et al.* (1957) have assumed that $T(r)$ is Gaussian. For metal foils, however, the heat conducted from the irradiated area is, within limits to be determined later, much larger than that radiated. So if the radiation is ignored and $H(r)$ is taken to be proportional to the experimentally determined $I(r)$, then the temperature will be given by

$$T(r) = T_0 + \frac{H_0 a^2}{4k}\left[\mathrm{Ei}\left\{-\left(\frac{r}{a}\right)^2\right\} - \mathrm{Ei}\left\{-\left(\frac{b}{a}\right)^2\right\} + \ln\left(\frac{b}{r}\right)^2\right] \qquad (2)$$

where T_0 is the temperature of the peripheral conductor of radius b, H_0 is the maximum heat flux and the Ei functions are standard exponential integrals (Jahnke and Emde 1945). For convenience the temperature across the foil is shown in Fig. 1 as the variation of the function $(4k/H_0a^2)\{T(r) - T_0\}$. The full and dotted lines correspond to a heat input described respectively by a Gaussian function and a pulse function of

height H_0 and width $2a$. It is evident that the maximum temperature T_m occurs at $r = 0$ and it is given by

$$T_m = T_0 + \frac{H_0 a^2}{4k}\left[\gamma + \ln\left(\frac{b}{a}\right)^2 - \mathrm{Ei}\left\{-\left(\frac{b}{a}\right)^2\right\}\right] \quad (3)$$

γ being Euler's constant. From this relation it would appear that the temperature becomes infinite when the boundary radius is infinite, independent of the area irradiated. However, this anomaly is removed in the complete time-dependent

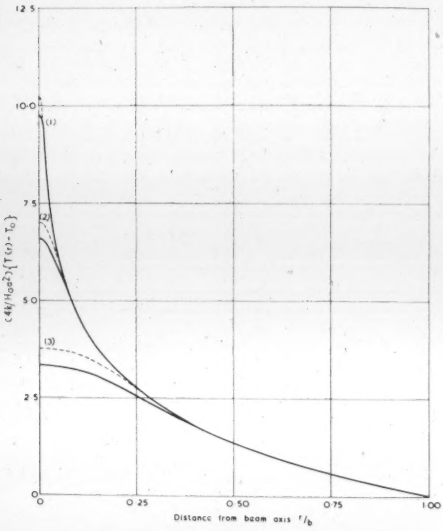


Fig. 1. Temperature distribution in the foils.
Curve 1, $a/b = 0{\cdot}01$; curve 2, $a/b = 0{\cdot}05$; curve 3, $a/b = 0{\cdot}25$.

solution, since for that case an equilibrium temperature distribution is not established.

The error involved in neglecting the radiation was estimated by calculating the ratio of the heat radiated H_R to the heat conducted H_C for a temperature distribution determined solely by the conductivity. Taking for convenience the pulse heat input function, the error will be given approximately by

$$\frac{H_R}{H_C} = \frac{\sigma\{1 + \ln(b/a)^2\}}{kd(T_m - T_0)} \int_0^b (T^4 - T_0^4) r dr. \quad (4)$$

This ratio decreases rapidly with decreasing b/a but does not depend sensitively on $T_m - T_0$. For extreme values, i.e. for an iron foil with $a = 10^{-3}$ cm, $b = 0{\cdot}5$ mm, $d = 10^{-5}$ cm, $T_m - T_0 = 1000$ deg C, $H_R/H_C \simeq 0{\cdot}1$. In general, however, $H_R/H_C = O(10^{-2})$ so that the conductivity approximation is justified.

In order to determine experimentally the relation between the maximum temperature and the size of the irradiated area, foils of iron, tin, copper and gold were mounted on the Siemens Stereoholder in the following ways: between a 200 mesh (70 μ) copper grid and the normal cooling ring; between two such grids superimposed; and using the cooling ring only. Observations on annealed iron foils about 3000 Å thick showed that the $\alpha-\gamma$ transformation, occurring at 910° C, could be produced consistently at the same beam diameter (17 μ), accelerating voltage (10^5 v), condenser aperture (600 μ) and beam current (50 μA) when they were mounted using either one grid or none. The transformation appears as a nucleation of regions having an orientation different to that of the matrix. The dislocation density in the matrix was observed to be very small, confirming that the sample was in the annealed condition, and the observed nucleation is therefore the α–γ transformation. The transformation could not be produced when the foil was mounted between two grids and this is evidently due to better thermal contact and hence greater heat conductivity. These results apply to regions well removed from the edge of the foils. For regions near the edge, the α–γ transformation could be produced with a beam diameter of only 11 μ (i.e. about half the heat input) for the one grid and no grid mounting conditions. This is presumably due to the asymmetrical heat conduction. Occasionally it was possible to melt the edges of iron foils, but this was never so with copper and gold which have much higher thermal conductivities. Less consistent results were obtained with the melting of tin foils (m.p. 232° C) when mounted on one grid, which occurred at beam diameters between 5 and 9 μ. This inconsistency was ascribed to non-uniform thickness and oxidation. Further observations made on iron–carbon foils showed that the solution of cementite lamellae took place when the beam diameter was between 15 and 17·4 μ. Considering all the results, it was decided to use the temperatures found for the pure iron foils to determine H_0. Using this value the curve represented by Eqn (2) is shown plotted in Fig. 2, together with experimental results obtained for solution of cementite and melting of tin. Since these results refer to specimens having effective thermal contact only with the

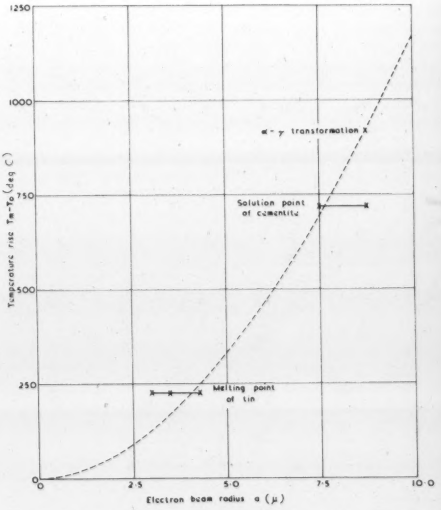


Fig. 2. ×, beam radius at which transformation or melting first occurs.

cooling ring, b was taken to be its inside radius: 0·50 mm. The value of H_0 is given also by the relation $H_0 = \epsilon I_0 v/d$, where v is the accelerating voltage and ϵ is the fraction of the electron energy absorbed. Comparing the two values of H_0, ϵ can be determined, and taking $d = 3 \times 10^{-5}$ cm and $v = 10^5$ v: $\epsilon \simeq 0{\cdot}10/I_0$. The value of ϵ can also be estimated from the Thomson–Whiddington law (Zworykin *et al.* 1945), which in these circumstances leads to a value of about $1{\cdot}2 \times 10^{-2}$. This implies that $I_0 \simeq 8{\cdot}0$ a/cm^2, a figure comparable with the data in the revised handbook for the instrument (Wolff, private communication).

A convenient minimum beam radius for most work with metal foils is $a \simeq 2\,\mu$. From Fig. 2 it is evident that the maximum rise in temperature of an iron foil when examined at 100 kv, 60 μa with a 600 μ condenser aperture is 70 deg c. This curve may also be applied to other metals by applying an appropriate thermal conductivity factor. Thus the maximum rise in temperature of a similar aluminium foil under these conditions would be about 30 deg c, and for gold about 20 deg c. Silcox and Whelan (1960) have deduced that under similar experimental conditions the temperature rise in aluminium is less than about 15 deg c.

From the temperature distributions and maximum temperatures it is now possible to calculate the stress distribution. For simplicity it was assumed that the foils did not buckle. Using normal elasticity methods (Timoshenko and Goodier 1951), the principal stresses were calculated for a Gaussian heat input and for two boundary conditions: the foil anchored at its periphery and not anchored. The values for these two conditions are similar. Taking the first more realistic case σ_{rr} and $\sigma_{\theta\theta}$ are given by

$$\sigma_{rr} = -\frac{\alpha E H_0 a^2}{8k}\left\{1 + \mathrm{Ei}\left\{-\left(\frac{r}{a}\right)^2\right\} - \mathrm{Ei}\left\{-\left(\frac{b}{a}\right)^2\right\} + \ln\left(\frac{b}{r}\right)^2 - \left(\frac{a}{r}\right)^2\left[1 - \exp\left\{-\left(\frac{r}{a}\right)^2\right\}\right] - \frac{1+\nu}{1-\nu}\left(1 - \frac{a^2}{b^2}\left[1 - \exp\left\{-\left(\frac{b}{a}\right)^2\right\}\right]\right)\right\} \quad (5)$$

$$\sigma_{\theta\theta} = -\frac{\alpha E H_0 a^2}{8k}\left\{-1 + \mathrm{Ei}\left\{-\left(\frac{r}{a}\right)^2\right\} - \mathrm{Ei}\left\{-\left(\frac{b}{a}\right)^2\right\} + \ln\left(\frac{b}{r}\right)^2 + \left(\frac{a}{r}\right)^2\left[1 - \exp\left\{-\left(\frac{r}{a}\right)^2\right\}\right] - \frac{1+\nu}{1-\nu}\left(1 - \frac{a^2}{b^2}\left[1 - \exp\left\{-\left(\frac{b}{a}\right)^2\right\}\right]\right)\right\} \quad (6)$$

where α is the coefficient of linear expansion, E is Young's modulus and ν is Poisson's ratio. The shear stresses S are $\frac{1}{2}\sigma_{rr}$, $\frac{1}{2}\sigma_{\theta\theta}$ and $\frac{1}{2}(\sigma_{rr} - \sigma_{\theta\theta})$ and curves of greatest shear stress were calculated for $a/b = 0{\cdot}01$, and $0{\cdot}05$ and $0{\cdot}25$ and are shown in Fig. 3. The maximum shear stress occurs when $r = 0$ and is given by $\frac{1}{2}\sigma_{rr} = \frac{1}{2}\sigma_{\theta\theta} = \frac{1}{4}\alpha E(T_m - T_0)$. Inserting the values of E and α for iron, with $T_m - T_0 = 70$ deg c, the maximum shear stress is about 4·5 kg/mm^2. It was found that under these conditions dislocations could rarely be moved, which is in agreement with the known value of the frictional shear stress of about 5 kg/mm^2 (Allen *et al.* 1956), since buckling would somewhat reduce the thermally induced stress. For aluminium examined under the same conditions the temperature rise would be about 30 deg c and the induced stress about 1·3 kg/mm^2. In this case the frictional stress including surface traction was found to be of the order $\frac{1}{2}$ kg/mm^2 and the electron beam would

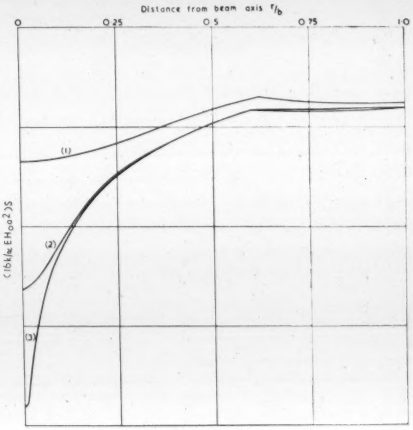


Fig. 3. Variation of maximum shear stress.
Curve 1, $a/b = 0{\cdot}01$; curve 2, $a/b = 0{\cdot}05$; curve 3, $a/b = 0{\cdot}25$.

be sufficiently intense to change the dislocation arrangements completely. From Eqn (2) it appears that $T_m - T_0$, and consequently the maximum shear stress, would be reduced by a factor of about 2 if the effective value of b is decreased from 5×10^{-2} to 3×10^{-3} mm by using two superimposed grids.

Acknowledgment

The work described above has been carried out as part of the general research programme of the National Physical Laboratory and this paper is published by permission of the Director of the Laboratory.

References

Allen, N. P., Hopkins, B. E., and McLennan, J. E., 1956, *Proc. Roy. Soc.* A, **234**, 221.

von Borries, B., and Glaser, W., 1943, *Kolloid Z.*, **106**, 123.

Carrington, W., Hale, K. F., and McLean, D., 1960, *Proc. Roy. Soc.* A, **259**, 203.

Jahnke, E., and Emde, F., 1945, *Tables of Functions* (New York: Dover Publications).

Silcox, J., and Whelan, M. J., 1960, *Phil. Mag.*, **5**, 1.

Timoshenko, S., and Goodier, J. N., 1951, *Theory of Elasticity* (New York: McGraw-Hill).

Whelan, M. J., 1957, *Dissertation* (University of Cambridge).

Whelan, M. J., Hirsch, P. B., Horne, R. W., and Bollmann, W., 1957, *Proc. Roy. Soc.* A, **240**, 524.

Zworykin, V. K. *et al.*, 1945, *Electron Optics and the Electron Microscope* (New York: John Wiley).

The frictional properties of lignum vitae

by K. G. McLAREN,* M.Sc., and D. TABOR, Sc.D., Research Laboratory for the Physics and Chemistry of Solids, Cavendish Laboratory, Cambridge

MS. received 10*th October* 1960

Abstract

A study has been made of the frictional properties of lignum vitae, an extremely hard wood which finds an important application in underwater bearings. In the dry state, the friction of this material is comparable with that of P.T.F.E. and experiments show that this is due to the lubricating action of woodwaxes expressed from the wood during sliding. The rolling friction due to hysteresis losses is very small. The sliding friction can be explained satisfactorily in terms of an adhesion mechanism.

Introduction

LIGNUM VITAE (the popular name for wood of the genus *Guaiacum*) is a very versatile material. Originally prized for the medicinal value of its extractives, its unusual frictional properties have been utilized for items as far apart as wooden clocks and stern tube bearings on ocean liners. As early as 1717 we find James Harrison, a clockmaker, introducing lignum vitae roller pinions 'which move so freely as never to need any Oyl' (see Lloyd 1953). In the mid-19th century, this wood was found to be the answer to the problem of providing underwater bearings for screw-driven ships, where metal corrosion ruled out metal-to-metal contact. For this purpose, it is still regarded as pre-eminent (Hide 1956), and gives years of wear (Jane 1955). Resins and waxes are distributed throughout the wood in special ducts (Fig. 1), and these give it a waxy feel and self-lubricating properties which are utilized in pulley-blocks, rollers and slides.

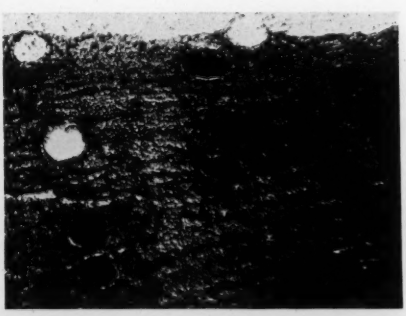


Fig. 1. Radial–transverse section of lignum vitae showing resin ducts. × 80.

Its 'built-in' lubrication, great hardness and durability make lignum vitae a very interesting natural product, and this paper describes a study of some of the factors involved in sliding and rolling friction on this material.

* Seconded from Australian Defence Scientific Service, Defence Standards Laboratories, N.S.W. Branch, Sydney.

Experimental

Test specimens of heartwood were sawn from a log cross section, then machined with a milling cutter to dimensions of 5, 3·5 and 1·5 cm in the longitudinal, tangential and radial directions, respectively. The finished surfaces were smooth and glossy. Friction measurements were made on longitudinal–tangential faces, by means of a slightly modified Eldredge apparatus (Eldredge and Tabor 1955), using steel balls as sliders or rollers. Traversing speeds were about 2 mm/sec.

Sliding and rolling friction of steel on dry, machined lignum vitae

The moisture content, taken as the ratio of the loss of weight at 60° C, to the oven-dry weight of the wood, was 6·8%. (Care was taken in the drying, since if the wood is heated at 100° C resinous material exudes freely, setting to a brittle mass on cooling.)

In the dry, machined condition, the coefficient of friction μ varied by as much as $\pm 15\%$ from point to point over the surface of specimens. Within this rather large range of experimental error, the friction was essentially independent of load (up to 4 kg) and ball diameter (5 mm to 20 mm).

Sliding friction μ_s ranged from 0·10 to 0·13, and rolling friction μ_r from 0·03 to 0·05, over a number of specimens. Faint permanent tracks were produced in the wood by both rolling and sliding. There was no significant change in the friction when repeated traverses were made along a given track. With heavy loads, a thin film of wax expressed from the wood was observed on the ball.

If we assume, as in the earlier work with balsam wood (Atack and Tabor 1958), that the sliding friction μ_s is composed of an adhesion term μ_a and a deformation term, the latter is very nearly the same as the deformation involved in rolling μ_r. We may therefore write

$$\mu_s = \mu_a + \mu_r.$$

Typical results for a 5 mm ball are given in Fig. 2. It is seen that μ_a is roughly independent of load and has a value $\mu_a \simeq 0{\cdot}08$.

Sliding friction of wet lignum vitae

Specimens soaked in water for eleven days at room temperature had a distinctly slippery feel, which was removed by a final wash in running water. A certain amount of material was leached out of the wood, but water uptake was small (about 2% on a dry basis as before). A thin film of water was present on the surface while measurements were being made. The results were not affected if this film was wiped off before a friction run. The water immersion treatment increased the sliding friction significantly only at low load. For a 320 g load the average value of μ_s was $0{\cdot}19 \pm 0{\cdot}03$, but for 3200 g load, μ_s was $0{\cdot}14 \pm 0{\cdot}02$, near the original (dry) value.

The effect of extraction of resins and waxes

To facilitate complete extraction, specimens were reduced to $\frac{1}{4}$ in. in thickness, then extracted successively with acetone, petroleum ether, benzene, alcohol and water in a Soxhlet apparatus. The extracted specimens were finally abraded under running water on 500 grade silicon carbide paper.

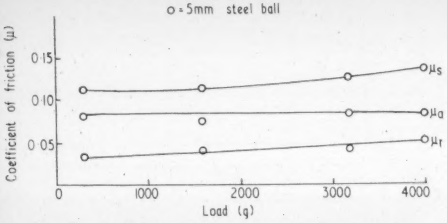


Fig. 2. Coefficient of friction of steel on lignum vitae.

The extraction process leads to an increase in the sliding friction. There were considerable variations in the friction along a given track, and from one track to another. In the saturated condition, for both 5 mm and 20 mm balls, μ_s was in the range 0·3–0·45 at low loads (320 g) and 0·2–0·3 for a load of 3200 g. The motion was intermittent, but there were no readily discernible permanent tracks left in the wood. As the wood was allowed to dry out at room temperature in a desiccator, sliding became smoother, but tracks were visible after sliding or rolling.

Extraction of the wood had relatively little effect on rolling friction, μ_r ranging from 0·03 to 0·08 for both wet and dry specimens.

The lubricating effect of the wood waxes

The products recovered after the extraction processes were all partly oxidized and had rather a sticky consistency. A short extraction of lignum vitae sawdust with cold petroleum ether gave a whitish waxy substance, which was found to act as a lubricant when rubbed into dry extracted wood. Under the extreme load conditions used (3200 g on a 5 mm ball), this 'wax' reduced μ_s in a typical case from 0·24 to 0·10, i.e. to the original value for the unextracted wood.

The effects of extraction and subsequent treatment on the frictional behaviour are shown in Fig. 3.

The effect of temperature on the friction of lignum vitae

Thin specimens were heated on a metal plate, the surface temperature near the slider being measured by a chromel–alumel thermocouple. A 5 mm ball was used, and a normal load of 800 g. There was a slight decrease in μ_s up to 80° C. At higher temperatures, gum and resin bubbled out of the wood, and results became very erratic.

Friction of lignum vitae on steel

In order to eliminate the deformation term in the frictional process, experiments were carried out in which hemispherically shaped wooden sliders of various radii were slin over a flat steel surface, previously abraded on wet silicod carbide papers, then degreased with acetone and petroleum ether. Under these conditions, after the first traverse, the deformation factor would be negligible, and the friction would be mainly due to adhesive forces. The coefficient of sliding friction was of the same order (0·14 ± 0·02) as that found for steel sliders on the wood, and was independent of load and radius of curvature of the slider. In certain cases (with a low load (320 g) and prolonged traversals over fresh surfaces of steel), the friction gradually built up and values of μ_s up to 0·5 were recorded. This was apparently due to local exhaustion of the lubricating wax at the surface of the wood. With heavier loads (3200 g) this effect was not observed, presumably because there was a greater amount of wax expressed from the wood.

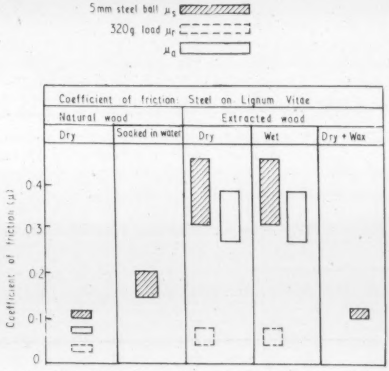


Fig. 3. The effect of wood-waxes on the friction of lignum vitae.

The effect of extraction on the hardness of lignum vitae

Indentations were made by a 2 mm steel ball mounted in a standard Vickers Hardness machine. The impressions were elliptical, and the radius of the circle with area equivalent to the ellipse was calculated.

The average radius varied with load according to a power law, the index lying between 2·3 and 3. This indicates that the deformation is neither purely plastic (index = 2) nor elastic (index = 3), but an intermediate type of viscoelastic deformation. The mean indentation pressure or hardness is thus dependent on load, but for simplicity we describe results obtained at a load of 10 kg. The hardness of the natural wood was about 23 kg/mm^2 and of the extracted wood about 15 kg/mm^2. For comparison some typical indentation hardness values for other materials are given in the Table. As a matter of interest typical friction values largely taken from Bowden and Tabor (1954), are also included.

It is seen that lignum vitae is about 40 times harder than

Approximate indentation hardness and some representative friction values

Material	*Indentation hardness (kg/mm²)*	*Coefficient of friction (Steel on material)*
Lignum vitae, natural	23	0·1
Lignum vitae, wax-extracted	15	0·4
Balsam wood, dry	>0·5	0·6
Balsam wood, wet	0·5	0·5
Polythene (polyethylene)	2	0·4
P.T.F.E. (polytetrafluoro-ethylene)	5	0·1
Nylon	10	0·3
Polystyrene	20	0·45
'Perspex'	20	0·45
White metal bearing alloys	20	0·4
Copper–lead bearing alloys	40–60	0·2

balsam wood and comparable in hardness with the harder polymers and with typical white-metal bearing alloys.

Discussion

The experiments show that the friction of dry lignum vitae in its natural state is quite low, and is comparable to values observed with P.T.F.E. Greene (1959) has recently reported that, under different experimental conditions, lignum vitae and nylon sliding on steel have the same coefficient of friction (0·16). The low friction is due, as the extraction experiments show, to a thin film of wood-wax either present on the surface or extruded by the sliding process itself. The conditions are virtually those of boundary lubrication, and the low coefficient of friction observed is almost independent of load. Since the wax is present as a three-dimensional supply, the surface layer is readily replenished and the friction does not increase markedly with repeated traversals of the same track.

If wood-waxes are extracted, the friction increases to values as high as $\mu = 0{\cdot}45$. Extraction leads to a slight softening of the wood (a decrease in indentation hardness of about 30%), but this is not accompanied by an appreciable increase in the deformation losses involved during rolling (μ_r) or sliding. The increase in sliding friction must thus be due primarily to increased adhesion between the slider and the extracted wood. Indeed, the friction is of the same order as that observed with other polymeric solids where marked adhesion is known to occur (see Table, column 3). If the wood-wax is re-applied to the extracted wood, the friction falls to approximately its original value.

These results show that the friction of lignum vitae can be explained in a fairly satisfactory way in terms of an adhesion mechanism. The low friction generally observed is primarily due to the lubricating action of the wood-waxes expressed from the wood during sliding.

Acknowledgments

We thank Dr. F. P. Bowden for his interest in this work, Messrs. Small and Parkes for a research grant, Captain P. H. Craven-Phillip of the Admiralty Engineering Laboratory who supplied the lignum vitae and Vauxhall Motors for a grant to the Laboratory.

References

ATACK, D., and TABOR, D., 1958, *Proc. Roy. Soc.* A., **246,** 539.

BOWDEN, F. P., and TABOR, D., 1954, *The Friction and Lubrication of Solids* (Oxford: Clarendon Press).

ELDREDGE, K. R., and TABOR, D., 1955, *Proc. Roy. Soc.* A., **229,** 181.

GREENE, S., 1959, *Forest Products J.*, p. 303.

HIDE, W. T., 1956, *Shipbuilder*, **63,** 644.

JANE, F. W., 1955, *The Structure of Wood* (London: A. and C. Black), p. 262.

LLOYD, H. A., 1953, *Suisse Horlog., October.*

Barium–carbon monoxide system: an hypothesis of Arizumi and Kotani

by T. A. G. GIORGI, M.Sc., Grad.Inst.P., and S. ORIGLIO, S.A.E.S. Getters Research Laboratory, Milan, Italy

M.S. received 5th September 1960

Abstract

According to Arizumi and Kotani carbon dioxide may be formed from carbon monoxide owing to the catalytic action of a barium getter deposit. This hypothesis has been experimentally investigated using a ion resonance mass spectrometer of the omegatron type. It has been shown, using the dynamic sorption measuring technique at constant pressure on the getter film, that there is no such catalytic action. The results obtained indicate that the experimental behaviour observed by Arizumi and Kotani, and which warranted their hypothesis, is due to outgassing of the sealed-off system.

Introduction

THE importance of the work carried out by Arizumi and Kotani in 1952, concerning carbon monoxide sorption by barium films, has been emphasized elsewhere (Ricca and della Porta 1961, della Porta and Ricca 1960). The work of Arizumi and Kotani is particularly interesting owing to the fact that they recognized the complexity of the phenomenon and underlined the importance played by film structure and by diffusion processes.

A number of the proposals made by Arizumi and Kotani have been experimentally verified. However, their hypothesis concerning the formation of carbon dioxide from carbon monoxide due to the catalytic action of barium has not, as yet, received such an experimental verification. On this hypothesis, which appears rather doubtful even at first sight, since it seems inconceivable that the carbon dioxide thus formed is not strongly sorbed by the getter film, is partially based their theory relating to carbon monoxide sorption at room temperature.

The experimental basis for this hypothesis was given by the changes in pressure which occurred in a closed vessel containing the getter film and in which a quantity of carbon monoxide had been introduced. Arizumi and Kotani had found that, under some circumstances, the pressure initially suffered a rapid decrease which was subsequently followed by a small but noticeable increase. They explained this phenomenon by assuming that carbon dioxide was released within the system owing to the catalytic action of barium on the introduced carbon monoxide.

Experimental

An experimental verification of this hypothesis may now be furnished by employing an ion resonance mass spectro-

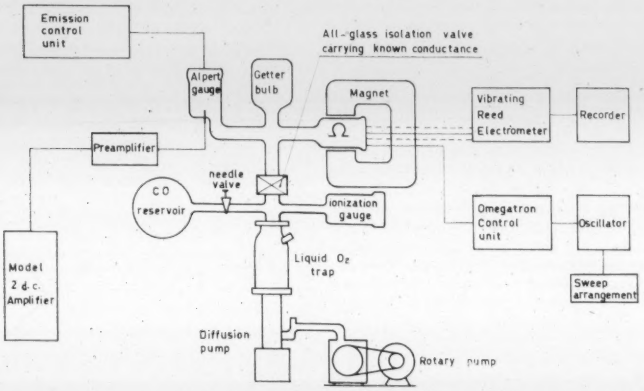


Fig. 1. System employed to study the gases above a barium getter working at a constant pressure.

meter of the omegatron type (Sommer, Thomas and Hipple 1951).

To this end the apparatus shown schematically in Fig. 1 was employed. It is essentially the same arrangement normally employed in this laboratory to study the gettering by barium films at constant pressure on the getter deposit.

The essential difference in this apparatus, besides the use of an omegatron head, rests on the fact that the capillary of known conductance, which is introduced between the getter chamber and the manifold, is carried by a grease-free all-glass spherical joint isolation valve (Yarwood 1957). Due to the use of this valve, and of an aluminium gasket (Holden, Holland and Laurenson 1959) to connect the manifold to the liquid oxygen trap, the initial residual pressure in the system is always better than 5×10^{-8} torr. Analysis at this pressure with the omegatron showed carbon monoxide to be the principal residual gas.

The getters used in these experiments are of Italian production (S.A.E.S. Getters, $St_2/7 \times 7$) and employ a stabilized barium aluminium alloy, from which pure barium is evaporated. The carbon monoxide used is produced by the method of Weinhouse (1948); any carbon dioxide which is also evolved is removed by allowing the gas to flow through three liquid oxygen traps in series.

After evaporation of the getter the capillary was introduced between the getter film and the pumps. The pressure on the getter was brought to $1{\cdot}5 \times 10^{-6}$ torr and maintained at this value until the initial sorption velocity ($6{\cdot}5 \times 10^{-12}$ mol. s^{-1}) had fallen to one-hundredth of this value. In this manner the various phases of sorption were followed and, in particular, the high initial velocity phase and the final phase characterized by a low sorption velocity. This latter phase is of particular interest, since it is to this that the work of Arizumi and Kotani refers.

Throughout these stages the mass region close to that of carbon dioxide (44 a.m.u.) was closely observed. Under no circumstances was it possible to observe a peak corresponding to the presence of carbon dioxide. Had its abundance been greater than about 0·5% of the principal carbon monoxide peak, its presence would have been easily revealed.

Owing to this it was decided to control the experimental basis of the hypothesis of Arizumi and Kotani. This was done by attaching a small bulb containing the getter to an Alpert type ionization gauge. The total volume of this system was of the order of 300 cm^3. To measure the pressure, the emission of the Alpert gauge was maintained constant at 0·01 mA. The whole was attached, by means of a constricted tubulation to facilitate seal-off, to an ordinary pumping system capable of reaching a residual vacuum of the order of 1×10^{-7} torr.

Two series of experiments were carried out using this type of system:

(*a*) the whole system was well degassed in order to obtain a residual vacuum of the order of 1×10^{-7} torr in the experimental tube;

(*b*) the whole system was only lightly degassed and for a brief period, so that the residual vacuum in the system was of the order of 5×10^{-5} torr.

In both cases after the getter deposit had been produced, carbon monoxide was introduced at a pressure of about 1×10^{-3} torr, and the getter bulb and ionization gauge were sealed-off from the pumps by glass-blowing.

The subsequent variation of pressure was recorded on a potentiometric recorder. A typical result for tubes treated in manner (*a*) is shown in Fig. 2, whilst Fig. 3 shows the resulting curve for tubes treated in manner (*b*).

It will be observed that the pressure increased, after the initial decrease, only if the sealed-off system had not been well degassed. This result was consistent for a whole series of experiments. It may therefore be deduced that the increase in pressure found by Arizumi and Kotani is not due to the formation of carbon dioxide, but rather to gas emission by the insufficiently degassed sealed-off system.

Further evidence to this effect was obtained by analysis of the residual gases above a barium getter, isolated from the pumping system by means of a spherical ground glass isolation valve, and on which a relatively high pressure of carbon monoxide had been established prior to seal-off.

Unfortunately, with the system only lightly degassed, it was found impossible to obtain a pressure below 1×10^{-5} torr in the experimental system when sealed off. Hence the omegatron could not be employed under these conditions, since at such a pressure an upper limit for its operation exists. Therefore the analysis had to be carried out in a well-degassed

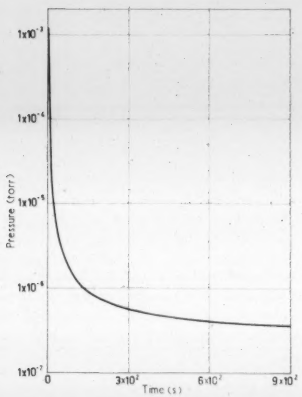


Fig. 2. Total pressure above a getter sealed off at a carbon monoxide pressure of 1×10^{-3} torr. System well degassed.

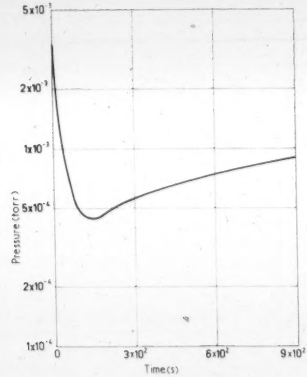


Fig. 3. Total pressure above a getter sealed off at a carbon monoxide pressure of 1×10^{-3} torr. System lightly degassed.

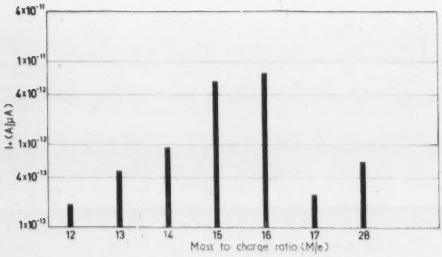


Fig. 4. Residual gases above a barium getter sealed off at a carbon monoxide pressure of 1×10^{-3} torr.
Total pressure $= 1 \times 10^{-6}$ torr; trapping = 0·2 v; r.f. voltage = 1650 mv.

system where, apparently, no emission of gas is perceivable, at least from curves of the type shown in Fig. 2.

In fact, the gases present on such a getter film, when immediately before seal-off the carbon monoxide pressure in the system was 1×10^{-3} torr, are shown in Fig. 4. This spectrum was obtained at a total pressure of 1×10^{-6} torr when the fall in pressure had become very slow. Although a great quantity of carbon monoxide was initially present, it can be seen that at this stage the peak at 28 a.m.u., where carbon monoxide and nitrogen appear together, is very small. No peak at mass 44 is observed. The main residuals are, in fact, the lighter hydrocarbons, particularly methane, which originate from heavier hydrocarbons owing to cracking on the hot filaments. Hence the gaseous phase at quasi-static equilibrium above a getter film isolated from the pumps consists of hydrocarbons which are released from walls of the enclosure and the electrodes.

Therefore, it seems likely that in a lightly-degassed system the getter film will initially sorb all the carbon monoxide introduced quite rapidly. However, in this case outgassing of hydrocarbons occurs in appreciable quantities. These are not quickly sorbed by the getter film and so an increase in the total pressure is observed after the initial decrease. After some time this increase tends to saturation and it has been found that a slow subsequent decrease in pressure occurs for very prolonged experiments. This is in all probability due to the similarity which exists between the ionization gauge-getter film combination and a getter ion pump. This effect could not have been observed by Arizumi and Kotani, since they employed a Pirani gauge for their pressure measurements.

Conclusions

It has been shown that there is no catalytic formation fo carbon dioxide due to the action of barium on carbon monoxide. However, the experimental results of Arizumi and Kotani have been shown to be correct. The interpretation then given by them is incorrect in as much as the gases evolved are hydrocarbons, which are released from the surfaces of the system when this is not well degassed. Hence their proposed theory for the adsorption of carbon monoxide at room temperature, although very illuminating on many points, seems to have been superseded by more recent theories based on experimental indications obtained employing far more refined techniques than then available, and which have been made possible by new developments in the field of vacuum technology.

Acknowledgments

The authors wish to thank Dr. P. della Porta, Managing Director of S.A.E.S. Getters, for permission to publish this paper.

References

ARIZUMI, T., and KOTANI, S., 1952, *J. Phys. Soc. Japan*, **7**, 415, 422.

HOLDEN, J., HOLLAND, L., and LAURENSON, L., 1959, *J. Sci. Instrum.*, **36**, 281.

DELLA PORTA, P., and RICCA, F., 1960, *Vide*, **85**, 3.

RICCA, F., and DELLA PORTA, P., 1960, *Vacuum*, **10**, 215.

SOMMER, H., THOMAS, H. A., and HIPPLE, J. A., 1951, *Phys. Rev.*, **82**, 697.

WEINHOUSE, S., 1948, *J. Amer. Chem. Soc.*, **70**, 442.

YARWOOD, J., 1957, *J. Sci. Instrum.*, **34**, 297.

Notes and comments

A history of flow measurement by pressure-difference devices

We have received a publication, under the above title, which consists of a reprinting of a series of articles, the first of which appeared in the very first number of *Instrument Engineer* in April 1952, and the last in October 1959. It consists of seven parts under the headings: The work of Venturi, The part played by Clemens Herschel, The development of the Venturi meter and some notable installations (by W. G. Ardley, formerly Joint Managing Director, George Kent Ltd.). The development of Venturi recorders from 1894 to 1911, The commercial metering of air, gas and steam (by John Lawrence Hodgson, B.Sc., Assoc.M.Inst.C.E.), Further diaphragm-operated instruments, including steam meters (by ffolliot Gray, formerly Development Engineer, George Kent Ltd.), Developments in differential producers for flow measurement (by H. E. Dall, Research and Development Department, George Kent Ltd.). It also contains nine appendices which give the theory of various instruments mentioned.

This publication does not claim to 'do more than scratch the surface of so wide a subject', but it should certainly prove of interest to those whose work brings them into the 'measurement field'.

It is published by George Kent, Ltd., and the price is 10*s*.

S.I.T./B.P.A. one-day conference

The Society of Instrument Technology and The Birmingham Productivity Association are holding a One-Day Conference, at The Birmingham College of Technology, Gosta Green, Birmingham, on Wednesday 29th March 1961.

The main purpose of this Society is to promote greater knowledge of instruments by the presentation of lectures and a series of papers upon the application of instruments and control to industrial processes and manufacturing units will be presented by engineers experienced in this work. Questions will be invited, and several histories of the improvements resulting from a more scientific system of measurement and control will be recounted and these histories related to the production problems which were resolved.

The lectures will stress (*a*) increased productivity using the same labour force, (*b*) consistent quality of products by controlled measurement, (*c*) reduction of scrap, during processing, by in-line inspection, including non-destructive testing.

An exhibition of the simple types of equipment which can effect improvements in production will be held and this will be closely related to the papers presented and the examples of installations which are recounted.

Convention on television and film techniques

A convention on television and film techniques, jointly organized by The Television Society and The British Kinematograph Society, will be held at The Institution of Electrical Engineers, Savoy Place, London, W.C.2, on 20th and 21st April 1961.

The convention will present some distinguished speakers both from Great Britain and from Overseas. A full programme of demonstrations is also planned. A noteworthy feature of the convention is that a considerably longer period than is normal has been set aside for discussion following each lecture, it being realized that delegates will welcome this innovation.

Full details of the convention and registration forms are available from: The Convention Secretary, The Television Society, 166 Shaftesbury Avenue, London, W.C.2. The registration fee for the convention is three guineas which includes the cost of a copy of the Convention Proceedings and discussion.

Second conference of x-ray analytical methods

A conference will be held on x-ray analytical methods in conjunction with the Department of Geology at The University, Manchester, from Monday 20th March to Friday 24th March 1961.

This year the conference will discuss and evaluate the latest techniques and instrumentation of x-ray fluorescence analysis. An extensive range of modern equipment will be available for demonstrating the principles of these methods over a wide field of applications.

Full information concerning the conference together with details of the registration fees may be obtained from the Conference Secretary at Research & Control Instruments Ltd., Instrument House, 207 King's Cross Road, London, W.C.1, to whom all inquiries and applications should be addressed.

Second International Conference on Materials Handling

The Second International Conference organized by the Institute of Materials Handling is to take place in Southport from May 10th to 12th at a time of particular importance for British production.

The conference will be the second devoted to materials handling and it was because the first, held in May 1959, was such a success that the Institute has decided to hold another.

The conference will provide a forum for the international exchange of knowledge and experience and the opportunity to discuss new methods and trends for the future in a field vital to delegates of all countries whose participation will give the conference its international flavour.

Brochures may be obtained from the Institute of Materials Handling, 32 Watling Street, London, E.C.4.

Euratom, U.K.A.E.A. and U.S.A.E.C. set up translation pool

The European Atomic Energy Community (Euratom), the United Kingdom Atomic Energy Authority and the United States Atomic Energy Commission have agreed to pool their efforts to collect and disseminate information concerning translations of nuclear literature, especially from such languages unfamiliar to Western readers as Russian and Japanese. A central information office 'Transatom' has therefore been established at Euratom's Brussels headquarters. It will function in two ways: first, by publishing

from December 1960 onwards a monthly *Transatom Bulletin*, which will list existing translations recently reported to the Brussels office as well as new translations planned by international or national institutions and private firms in the European Community, the United States, the United Kingdom and in other countries.

Secondly, all data relating to translations, including those made before the establishment of Transatom, are being collected and recorded in a master file system at Brussels. Copies of this card file have been offered to appropriate institutions in countries with great interest in the nuclear field.

Contacts have been established in order to avoid duplication of work when the European translation centre, to be established at Delft, Holland, is set up. The scope of this institution is much wider: it will cover all scientific and technical material in the field of exact sciences.

The *Bulletin* is available on a subscription basis from Transatom, 51 Rue Belliard, Brussels, Belgium, at $8 per annum, $16 air mail.

Forthcoming publication of the proceedings of the 1960 PICC Symposium, Rome

The *Symposium on the numerical treatment of ordinary differential equations, integral and integro-differential equations* took place in Rome during the week of 20th–24th September 1960 at the Mathematical Institute of the University of Rome. This symposium was organized by the Provisional International Computation Centre (PICC).

The symposium opened with a remarkable report delivered by Professor Walther of Darmstadt (Germany) on the methods presently employed in the treatment of integral and integro-differential equations. Dr. Genuys (Paris) then presented a second report, on the methods of treating ordinary differential equations. After this introduction, more than fifty specialists, divided into three study groups and lectured on their personal studies, exposing the particular problems with which they dealt and how the practical and theoretical difficulties which they encountered had been overcome.

Professor Lanczos (Dublin) delivered a more general lecture on the possibilities offered by modern electronic computers and the final session was consecrated to an outstanding speech by Professor R. Courant of New York, who exposed his personal conclusion on the requirements of scientific research in our highly technical century, as well as the problems posed by the training of young specialists in the field of automatic computation.

The symposium, on the whole, presented a fairly complete picture of the actual state of this important section of mathematical sciences. It was attended by about 200 eminent mathematicians from twenty different countries.

The *Proceedings*, consisting of about 700 pages of high scientific value, will be published by Birkhauser Verlag (Basel/Stuttgart) at the beginning of 1961.

Editorial and publications sales offices

Readers are requested to note that all correspondence of an editorial nature should now be sent to the following address:

The Editor and Deputy Secretary,
The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7 (Kensington 0048).

Correspondence concerning subscriptions, missing journals, and sales of publications should be sent to:

The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1 (Belgravia 6111).

Journal of Scientific Instruments

Contents of the March issue

PAPERS

Technique for measuring the elastic properties of bitumens, tars and soils under dynamic loading. By E. N. Thrower.
Simple amplifier and pulse-height analyser for an X-ray scintillation counter. By B. B. Trott.
Some results obtained in developing a recording Geiger-counter X-ray intensity monitor. By A. E. Kiss and N. Patla.
Small high-temperature vacuum furnace. By G. Arthur and J. H. Priest.
Calculator for solving electron irradiation problems. By E. H. Cornish.
Very high speed dual-purpose image-converter camera. By F. H. Grover.
Flat-field microscope objective. By C. G. Wynne.
Uni-axial pantograph. By J. C. Barnes.
A differential absorption spectrometer for the infra-red. By M. D. Sturge.
Analogue method for computing the number of spins contributing to a magnetic resonance absorption. By V. R. Burgess.
Servo-operated respiratory waveform simulator. By D. W. Hill, J. R. Hook and E. G. Bell.

LABORATORY AND WORKSHOP NOTES

The outgassing of large vacuum-systems to a temperature of 450° C. By A. L. Boers and J. B. van Dien.
A simple method of monitoring the rate of deposition of a conducting film in a vacuum chamber. By K. Feldmesser.
A clip for the rapid attachment of Warburg flasks. By R. J. Romani.
Verification of low loads in hardness testing machines. By R. J. Ellis.

NOTES AND NEWS

New instruments, materials and tools Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique).

All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.

REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611–2.)
SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12*s.* 6*d.* each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.
CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

UNWIN BROTHERS LIMITED, WOKING AND LONDON

The Joint Conference of the Non-Destructive Testing Group and the Société Française de Métallurgie

General report of the Conference

Abstract

This paper consists of a general report of the Joint Conference of The Non-Destructive Testing Group and the Société Française de Métallurgie, held in the Lecture Theatre of the Institution of Mechanical Engineers from 2nd to 4th May 1960. It includes a list of all the papers presented at the Conference, with references, where available, for those which have been or are being published, and is followed by four of the papers presented at the Conference.

THIS Joint Conference of the Non-Destructive Testing Group of the Institute of Physics and The Physical Society and the Société Française de Métallurgie on 'Physical and Structural Aspects of Non-Destructive Testing and Metals', was planned to include research methods or techniques from which new methods of physical testing could eventually develop.

Some of the papers presented are published with this summary; others, it is hoped, will be published in metallurgical journals either in France or in this country. The introductory paper 'The Structure of Non-Destructive Testing' by K. W. Andrews (this issue p. 127) explained the scope of non-destructive testing and physical methods in general, and the relationships between physical properties and structure with particular reference to metallurgical materials.

In the first session, Relaxation and Stress Wave Phenomena, three papers illustrated the use of physical techniques in which the specimen is subjected to elastic vibrations, and the absorption of energy at critical frequencies (temperature dependent) is measured and related to an atomic diffusion process. The determination of the damping of oscillations at different frequencies and temperatures[1]* could be used to study grain boundaries in metals, and the effects of both interstitial and substitutional elements at grain boundaries and the relation to embrittlement and grain boundary mobility could be determined. Hydrogen in steels could be studied[2] by a related technique employing an electrostatic apparatus using the resonant frequency of a specimen (10 000, 30 000 and 100 000 c/s according to size of specimen). This suggested that hydrogen was not in interstitial solution below about 150° C but in fact preferred dislocations or other sites. On the other hand excess hydrogen could be forced into the material, when it would show an interstitial damping peak at low temperatures. The significance of these results was indicated. A third method, the torsional pendulum method,[3] enabled the investigation of the precipitation of nitrogen in iron, and of copper of varying purity.

On the other hand, Dr. Wilks[4] gave a general and theoretical account of the movement of dislocations as such (i.e. not atoms or ions) under the influence of an applied oscillating stress. These movements of the dislocations at stresses much smaller than the critical shear stress also give rise to damping. Experimental and theoretical work on these lines is fully consistent with the general viewpoint of the Conference and the interpretation suggested by Figs 1 and 3 of the introductory paper by Andrews.

The final paper in this session, on the application of forced vibrations,[5] deserves separate mention for its closer connection with the classical physics of vibrations (i.e. where there is no immediate structural interpretation in terms of, say, atoms or dislocations). The system was arranged so that two pendulums together gave a pulsating beat frequency which is transmitted to the specimen under test. This method of forced vibrations can be used to determine elastic modulus or damping capacity.

The second session was concerned with Magnetic and Electrical Properties of Metals. There was a discussion of the difficulties involved in attempting an adequate mathematical treatment of induction in non-uniform magnetic fields[6] (this issue p. 146). A consideration of a metallic conductor as an infinite number of closed circuits led to an expression for the current at any point which depended upon the geometry of the conductor, the electrical conductivity and the magnetic permeability. It was considered that application of this theory of coupled circuits should lead to a re-examination of methods of testing.

The 'electro-test' method[7] is a non-destructive one based on physicochemical phenomena. It is a method for the identification of metals and alloys, and can be used as a spot test. A continuous e.m.f. of a few volts is applied between the metal and an electrode, and characteristic reactions occur according to the elements present. The method is generally applicable to steels including stainless and low alloy steels and to non-ferrous alloys, but in some cases silicon and carbon could be tested.

Magnetic Examination of Metals and Alloys[8] (this issue p. 141) was of use in the study of metallurgical phenomena such as precipitation, and recent work had been done on the relation between magnetic properties and stresses including the stress fields of dislocations. Surface defects on hot rolled square billets made from a mild (rimming) steel could be detected by an electromagnetic method. This had an important industrial application[9] in the development and improvement of a method for the continuous examination of a product in the course of processing. The basic physics of the device was explained and the method had resulted among other things in improved sensitivity and better resolution lengthwise.

The third session was concerned with Metallographic and New Techniques. The technique of examining a fractured surface by electron microscopy had been applied[10] to the study of fracture in steels containing hydrogen. In the presence of hydrogen the number of nuclei from which the fracture propagates is considerably increased, and helps to confirm current views about the way in which hydrogen

* Numbers refer to the individual papers, listed at the end of the article.

causes fracture. Problems of both hot and cold fracture in large objects such as crankshafts and large steel vessels could be studied by a similar technique.[11] Minor constituents from fractured surfaces could be removed and examined and it was found[12] that some of these constituents had undoubtedly provided a starting point for fracture.

Consideration of these techniques illustrate the point that, whilst applied non-destructive testing aims at establishing the existence of causes of failure and assessing their possible seriousness, the interest for physicists lies in the physical phenomena involved. The metallographic approach provides another aspect—it is directly concerned with the basic structural causes. These give rise, either to physical effects which can be recorded or measured or to mechanical failure which has to be explained. The importance of structure is again endorsed.

Investigation of Surface Deformation dealt with somewhat different techniques and phenomena. This work[13] (this issue p. 134) showed that freshly deformed surfaces gave a photoelectric sensitivity to visible and near-ultra violet light. Examination of surface topography could be undertaken by means of electron microscopical techniques, by tracer techniques and by optical methods and examples of application were described. In this field a combination of a wide variety of techniques has provided much new information.

K. W. A.

(1) LEAK, G. M., F.Inst.P., *Grain Boundary Damping in Iron and Iron Alloys* (University of Manchester).

(2) HEWITT, J., *A Study of Hydrogen in Low Alloy Steels by Internal Friction Techniques* (United Steel Companies Limited, Rotherham).

(3) COLLETTE, G., *Note on Certain Transient Phenomena Arising from Internal Friction* (Institut de Recherches de la Sidérurgie).

(4) WILKS, J., *Dislocation Damping in Metals* (Clarendon Laboratory, Oxford).

(5) GIET and SORIN, MM., *A New Application of the Method of Forced Vibration to the Study of Complex Structures. Experimental Determination of Deflection Influence Lines* (Institut Supérieur des Matériaux et de la Construction Mécanique).

(6) GRANEAU, P., *Frequency Dependence of Induced Currents* (British Insulated Callenders' Cables Limited, London) 1961, *J. Electronics and Control*, in the press.

(7) GUITTON, L., *The Non-Destructive Identification of Metals and Alloys by the 'Electro-test'*.

(8) HOSELITZ, K., F.Inst.P., *Magnetic Examination of Metals and Alloys* (Mullard Research Laboratories, Redhill, Surrey).

(9) BEAUJARD, L., MONDOT, J., and DARSCH, C., *An Improved Electromagnetic Method for the Detection of Surface Defects in Hot-Rolled Square Billets made from Mild Rimming Steel* (Institut de Recherches de la Sidérurgie).

(10) AZOU, P., AZOU, MME, and BASTIEN, P., *The Configuration of Fractures in Dead-Mild Steel containing Hydrogen.*

(11) JACQUET, P. A., and MENCARELLI, E., *The Techniques of Non-Destructive Electron Micrography and Microfractography.*

(12) MESDAMES GAILLARD, FRANÇOISE, and WEILL, ADRIENNE, A.Inst.P., *Non-Destructive Extraction and Analysis of Minor Constituents present in Fracture Surfaces* (Ingénieurs Contractuels des Constructions et Armes Navales).

(13) GRUNBERG, L., SCOTT, D., and WRIGHT, K. H. R., A.Inst.P., *The Investigation of Surface Deformation* (National Engineering Laboratory).

The structure of non-destructive testing

by K. W. ANDREWS, D.Sc., D.Phil., F.I.M., F.Inst.P., United Steel Companies Limited, Research and Development Department, Rotherham, Yorks.

Introductory paper presented at the Joint Conference of the Non-Destructive Testing Group and the Société Française de Métallurgie, 2nd to 4th May 1960

Abstract

The physical basis of all non-destructive testing is emphasized. It is suggested that the 'structure' (in a general sense) of non-destructive testing is based on two principal forms of energy, viz. the spectrum of electromagnetic radiations and fields and the spectrum of particle movements or vibrations in solids. The two are involved together or separately and are interpretable in terms of the structure (in a more physical sense) of the material, i.e. the electrons, atoms, molecules, crystal structure, etc.

A unified scheme of relationships is suggested. Reference is made to a fundamental unity of the equilibrium physical properties of a perfect crystal so that physical methods of examination are similarly inter-related. The 'real' materials of industry, and in particular metals, can be regarded as derived from perfect single crystals by the introduction of increasingly serious forms of lattice 'imperfection' or irregularity such as solid solution strains, vacancies, domains, dislocations, etc., leading to the eventual appearance of grain boundaries, poly-phase structures, segregation, etc. The way in which various physical properties are used to interpret materials in those terms is indicated.

Introduction—Some definitions and criteria

THE subject of non-destructive testing has been approached from many different points of view and its precise scope has been the subject of considerable debate. We should not attempt any exact definition here but prefer to accept a wide if somewhat inexact interpretation. Some useful observations were made at a previous conference of the Group (Rait and Peiser 1957). We should aim (as these authors imply) to avoid the dangers of an interpretation which is either too wide or too narrow.

A statement made at the Chicago Conference in 1957 (van Ouwerkerk 1957) illustrates the narrowest viewpoint, viz. 'Non-destructive Testing as we know it today consists primarily of four methods—Radiography, Ultrasonics, Magnetic Particle Inspection and Penetrant Methods. These are the testing methods which have been *accepted by industry*' (our italics). The author expected, however, that these four methods would 'surely expand to a longer list which would help to answer the problems of industry'. His definition of magnetic methods which have been available for several and others have been aware of a number of electrical and magnetic methods which have been available for several years (Lewis 1951, Institute of Metals 1958a). Some of these methods are used for the continuous checking of products such as control of strip thickness by a radiation gauge or magnetic sorting methods to give but two examples. These are clearly within the scope of non-destructive testing and the present industrial interest is indicated for example, by the recent book of Hinsley (1959) and the encyclopaedic two-volume 'Handbook' of the American Society (McMaster 1959).

The activities of this Group take us somewhat farther. The primary concern is not with the technical details of the applications of testing methods, but with the underlying physical and structural aspects. A similar point of view is exemplified by the series of volumes entitled *Progress in Non-Destructive Testing* (Stanford and Fearon 1958, 1960).

We cannot under-rate the possible value of this approach which seeks to examine both basic properties and techniques which are already successful in the laboratory bearing in mind that they may lead to new developments of use to industry. The land on both sides must be explored and bridges built so there can be traffic to and fro. Furthermore, as at this conference, it is desirable not to be too precise as to where the legitimate boundaries of activity should lie. A broad view of the subject which extends a good long way in both directions is suggested by the wide scope of the papers at the recent Third International Conference on Non-Destructive Testing held at Osaka, Japan, in March 1960 (to be published).

The American handbook (p.4.7) points out that 'Any valid law of nature may serve as a basis for a useful non-destructive test if it provides reliable measurements that can be correlated with significant material properties, discontinuities or serviceability of the test objects'. It is then noted that tests utilize three types of 'probing media' which are based on

(1) Motion of matter.
(2) Transmission of energy.
(3) Inseparably combined motion of matter and transmission of energy.

As it happens a more fundamental approach may also be based on the supposition of two or three main divisions but of a rather different kind, recognizing that energy and matter —stationary or in motion—are involved. We suggest here that the general 'structure' of non-destructive testing is based on two intrinsic forms of energy represented by:

(1) The spectrum of electromagnetic radiations and electromagnetic fields
(2) The spectrum of particle movements or vibrations in solids.

These two are involved together or separately in the material with which we are concerned and so the third element in this conception is:

(3) The structure of the material.

Finally we imply an essential unity involving all three.

The word 'structure' is now employed in its more usual sense—it includes the electrons, atoms, molecules, crystal

structure, lattice energy and vibrations, etc. It is in terms of the *central* structural and particle concepts that we would like to think on the one hand of physical processes which involve radiations or, on the other hand, of the mechanical movements of atoms, ions, molecules, etc., to and fro in a solid which are imposed by mechanical or physical means or which are already present, and which are associated with the forms of atomic binding in the solid. Clearly we could stray from here into a consideration of chemical or metallic bonding theories or into other theoretical aspects which are obviously in some way related to our theme but too far away from the immediate objective. We do, however, wish to imply that structural concepts could be brought into the consideration of properties much more often than appears to be the case. It is indeed possible to use or interpret many methods of testing—mechanical or physical—and even to develop theories and yet to overlook the underlying structural aspects. Solid state physicists and metallurgists have, however, in recent years derived great stimulus from the concept of dislocations, and much that is known about mechanical properties is interpreted in terms of this or other essentially structural concepts (Cottrell 1953, Fisher *et al.* 1956). Some physical phenomena associated with non-destructive testing also require structural considerations on the part of those who develop equipment or do research (e.g. properties of transducers) but structure can otherwise be evaded!

Relation of structure to the two energy spectra

The diagram in Fig. 1 represents the viewpoint outlined above.

On the left-hand side we consider the electromagnetic spectrum—a familiar enough diagram by itself. The radiation emitted from alternating currents at relatively low frequency arises from the to and fro movement of electrons in conductors. We do not consider that there is a fundamental difference between a field surrounding a charged oscillating particle such as an electron, and the field which radiates from it (at the same frequency). They are represented by a set of relationships which include Maxwell's equations. Alternating current methods of testing and the induction processes using eddy currents therefore fit in. At 'zero frequency' we include electrostatic and magnetic (magnetostatic) fields.

It is well known that as we proceed to higher frequencies the process of generation of the electromagnetic wave becomes increasingly more associated with processes which involve transitions between energy states ($h\nu = \Delta E$). Throughout the range radiation can be scattered, absorbed, diffracted on passing through or coming into contact with matter.

Numerous methods of testing, laboratory examination, spectrochemical analysis employ the fields or radiations represented by this side of the diagram. Many of these are indicated in order to support the idea of a fundamental unity, although some would normally be regarded as outside the range of non-destructive testing.

The right-hand side of the diagram comprises methods which involve the application of stress—generally in a cyclic manner. The case of 'zero frequency' includes direct stress (hardness) measurements, etc., in so far as they involve only elastic deformations. The case of plastic deformation is omitted but will be mentioned later. The spectrum is thus generally concerned with elastic vibrations, such as are used at low frequencies for fatigue tests and then at higher frequencies for ultrasonic tests. The passage of sound through solids belongs to the same group of phenomena. Likewise the movements of atoms at thermal frequencies represent an upper limit from the point of view of what is called the vibrational analysis of materials. There is an apparent difference between vibrations put into the material by a test method and the thermal vibrations already present. The latter involve firstly the equilibrium positions of the atoms determined by a balance of attractive and repulsive forces and representing a minimum of potential energy as a function of distance. The thermal energy is then represented by vibrations about this potential minimum (including 'zero point energy'). Any other vibrations are superimposed and involve longer wavelengths. An important feature of all vibration experiments or tests is the proportion of imposed vibrational energy which becomes transferred to the thermal vibrations and appears as heat. The wavelengths of externally imposed vibrations can be very much longer than interatomic distances and the energies involved lower.

Implicit in any consideration of a vibration spectrum is the presence of the third element on the scheme. Nor can we go very far in the consideration of any of the electromagnetic phenomena as test methods without reference to the structure of the material itself. In the diagram the frequency scales are placed at the same levels. There is no wavelength scale for the vibration spectrum because the velocity of sound is not a constant. In the other case the wavelength scale obviously refers to waves in free space. Between these frequency scales we have indicated the main structural factors involved. There are no clear divisions along this 'band' (any more than on either side), but at the lower frequencies we are more concerned with the solid as a solid—composed of atoms, ions or molecules—in vibration—and situated on a crystal lattice. The electrical phenomena involve conduction electrons which have a somewhat separate significance here. As the electromagnetic radiation energy (or frequency) increases we progress to regions where the transitions between molecular vibrational and rotational states or other energy states (including magnetic resonance) become important. The next stage involves more particularly the electrons in orbits in atoms (or molecules)—the outer electrons responsible for optical spectra and then the inner electrons giving rise to x-ray spectra at even higher energies.

It will be noted that some frequencies on the one spectrum correspond to, or can be translated into effects on, the other —sometimes by the interposition of a suitable instrument or device. In general a solid with a given vibrational spectrum will be sensitive to electromagnetic radiation and may be able to absorb or emit radiation if energy transitions involved are comparable with and can be excited by the radiation ($h\nu \geqslant \Delta E$), e.g. thermal vibrations and infra-red radiations are approximately 'opposite'. The radiations of the highest frequencies on the electromagnetic side correspond to transitions inside the nuclei or result from the breakdowns and interactions of fundamental particles. The spectrum of particle vibrations or frequencies cannot reasonably be extended into this region. The present limit is represented by the 'atomic unit of frequency' which is the frequency represented by the energy of an electron in the ground state of a hydrogen atom (the frequency associated with the orbit according to the Bohr theory).

Other aspects of the diagram are evident, and need not be discussed.

Interrelation of properties

The existence of many interrelationships between physical properties which are usually treated as distinct is well known, but often ignored. We suppose that in theory it should be

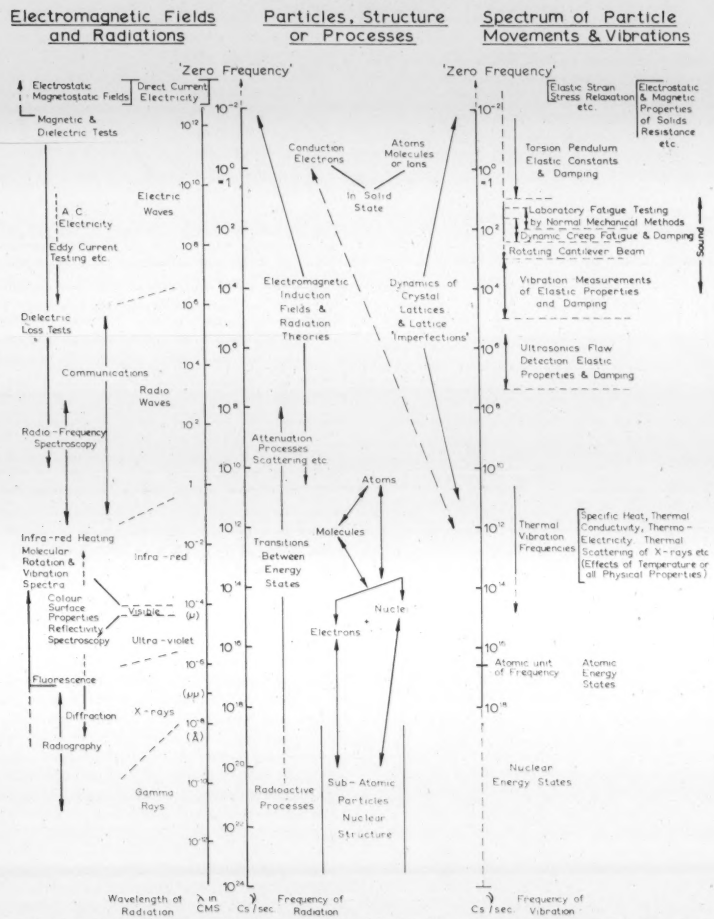


Fig. 1. General scheme of physical properties and phenomena used in non-destructive testing. Relationship between electromagnetic fields and radiation (left) or vibration spectrum (right) with atomic and crystalline structure (centre).

ultimately possible to calculate (perhaps by wave-mechanical methods) the complete energy relationships, electronic distributions, interatomic distances, etc., in solids. From the basic data the effects of external electric, magnetic or stress fields could then be established. In practice it is necessary to accept the limitations of both theory and experimental data. Some indication can, however, be given of the kind of relationship we have in mind, especially in regard to the relevant properties of solid materials.

Initially we consider an ideally perfect crystal. Some of its physical properties are governed by relationships which connect some kind of physical cause (externally applied) and an effect (the internal reaction). The simplest of these is typified by Hooke's law for an isotropic medium, stating that strain is proportional to stress.

A crystal is not isotropic and so we have the relationship between the tensor quantities:

$$\text{Stress} \xrightarrow{\text{elasticity}} \text{strain.}$$

Similar relationships exist between other quantities:

$$\text{Electric field} \xrightarrow{\text{permittivity}} \text{displacement.}$$

The strictly analogous relationship for thermal effects (zero order tensors) is represented by

$$\text{Temperature} \xrightarrow{\text{heat capacity}} \text{entropy.}$$

These relationships were summarized in the diagram, Fig. 2, due to Nye (1957). This diagram brings out physical

properties which involve not only the 'principal effects' such as the stress–strain relationship, but 'coupled effects' such as thermal expansion or the piezoelectric effect. Furthermore, the diagram has a certain symmetry, e.g. the coefficients for the piezocaloric effect are equal to the coefficients of thermal expansion and the coefficients represented by any side of the outer triangle are equal in magnitude, but opposite in sign to those of the parallel side of the other triangle. These effects can act together or separately.

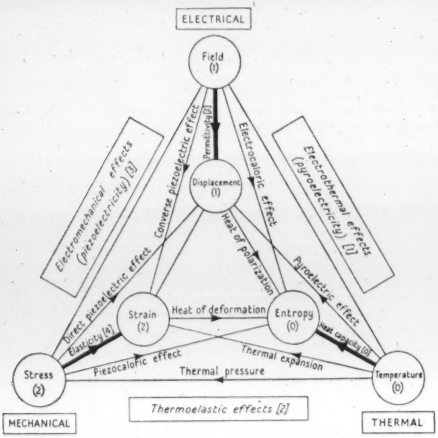


Fig. 2. Interrelations between certain (equilibrium) properties of a crystal.

Notes—(*a*) Three principal effects represented by heavy lines, coupled effects by the diagonal lines or triangle sides.
(*b*) The figures in round brackets are the tensor rank of the variables and in square brackets of the properties.
(*c*) All crystals can show the principal effects and some of the coupled effects (e.g. thermal expansion), but other coupled effects (notably piezoelectricity and pyroelectricity) can only occur if certain symmetry conditions are met.

The diagram is due to Nye (1957) and is reproduced with permission.

This interesting diagram contains a number of physical properties which either interest physicists concerned with non-destructive testing at present or could do so. It omits certain properties which we should now mention.

(i) An analogous triangle might be constructed with magnetic properties replacing the electrical. Alternatively there could be a tetrahedron with four principal effects, although the interrelations between magnetic and electric properties might not be strictly similar to the others. This possibility has been recognized by Nye (1957, p. 42, see also Cady 1946).

(ii) The relationships considered so far are first order effects: the tensor components of the properties are taken as constants. They are otherwise represented by partial differential coefficients. Second order variations are important for some phenomena which are (or could be) of practical use, for example,

(*a*) Photoelasticity is represented by small changes in certain tensor components which correspond to refractive index or dielectric constant, and involve a second order dependence on field ('electro-optical effect') and on stress ('piezo-optical effect').

(*b*) Variation of elastic constants with temperature.

(*c*) Second order variation of strain with electric displacement (electrostriction) or with magnetic induction (magnetostriction).

(*d*) Non-linear thermal expansion effects.

(*e*) Non-linear variation of magnetic permeability (ferromagnetism).

(iii) The treatment of optical effects is implicit in the relationships involving electrical properties and these represent one aspect of the interaction with electromagnetic radiation and with any electrical or magnetic field. Other important aspects require separate consideration (see below).

(iv) Electric currents in general apart from the so-called displacement current involve the flow of electrons along a potential gradient. The flow of heat along a thermal gradient is analogous and the relationship between the two (Wiedemann–Franz law) is well known. The other properties mentioned are indeed equilibrium properties, but these are transport properties and are not thermodynamically reversible. (The significance of the conduction electrons for both phenomena is represented in Fig. 1 by the broken line.)

Thermoelectric effects involve both types of conductivity simultaneously and the two processes interfere. In the most general case the heat evolved per unit volume at any point in a conductor carrying a current is the sum of three components: Joule heat + thermoelectric heat + heat conducted to the point.

From perfect crystals to real materials

The fundamental relationships should strictly apply only to a perfect lattice. Such a lattice can be regarded as departing progressively from the ideal state to a real material by the introduction of progressively more serious 'imperfections' or defects. The writer believes that the scheme in Fig. 3 represents something like the right order of 'progressive degeneration' from a perfect crystal to a real material with defects of the kind which non-destructive testing, as applied, aims to detect. The diagram (in keeping with the theme of this conference) has a bias towards the metallurgical materials.

We cannot consider in any detail how the physical properties for the ideal crystal are affected, but firstly we can summarize some of the main techniques or principles which apply and which need *not* be discussed further here:

(1) X-ray (and electron) diffraction methods provide information about solid solution effects and the identity of phases.

(2) Electron and optical microscopy provide information about particle size, morphology, etc.

(3) A number of investigation techniques are available for grain boundaries (McLean 1957, see also Weinberg 1959).

(4) Other methods are available for the study of magnetic domains (powder methods) (Bozorth 1951), and domains in ordered structures (x-ray diffraction).

(5) Low angle scattering methods are available for certain types of segregation or phase distribution (Guinier and Fournet 1955).

(6) Electron microscopy in general is being used for the study of dislocations, etc. Interesting information is becoming available from the thin film methods (Institute of Metals 1959).

(7) The effects of vacancies and other point defects on

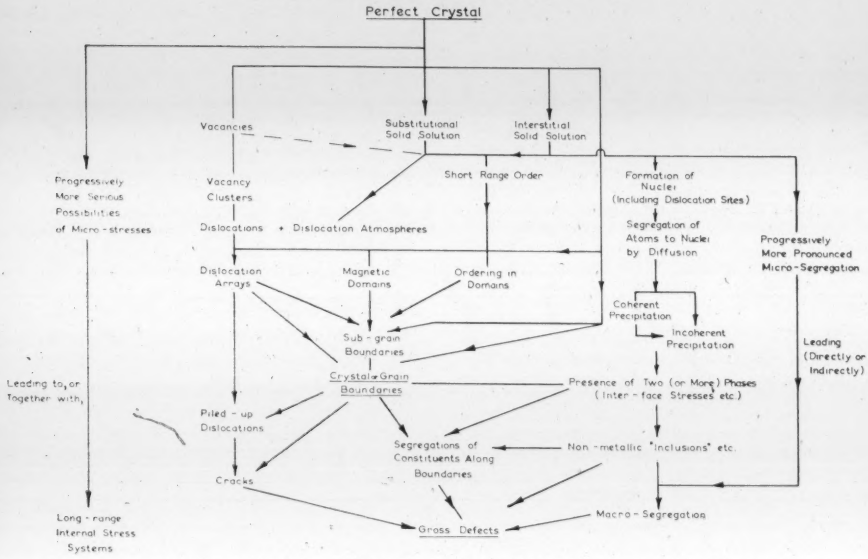


Fig. 3. Progressive degeneration from perfect crystal to real materials with increasing possibility of internal stresses, segregation or defectiveness.
(Particular relevance to metallurgical materials.)

the properties of metals, alloys or other solids have been described elsewhere (Institute of Metals 1958b).

(8) It is an important principle that, whereas single crystals are generally anisotropic, a polycrystalline material will tend to show pseudo-isotropic properties which average the properties for different directions of a single crystal. Varying degrees of preferred orientation affect the extent of anisotropy shown by a given piece of material. Inter-granular stress conditions will be more serious the more pronounced the anisotropy of the individual crystals. This could, for example, be an important factor affecting the properties of polycrystalline hexagonal metals.

(9) The effects of dislocations on the mechanical and other properties of metals have been the subject of extensive studies (Cottrell 1953, Fisher *et al.* 1956) and the concepts are now well known and frequently used by physical and structural metallurgists.

Apart from these general observations we require some more specific indication as to how the physical properties mentioned in the previous section are affected.

In the first instance elastic vibrational behaviour derives from an alternating stress or force of some kind. The energy of the elastic waves can be lost by several processes. A number of these processes have been examined by Mason (1958). We are here interested in such effects as

(1) Loss of energy by thermal conduction (longitudinal waves only) which, for example, in polycrystalline materials leads to losses between grains due to the anisotropy and change in orientation at boundaries (the Zener effect (Zener 1948)).

(2) Loss due to scattering when wavelength is of the same order as the grain size. These could also be associated effects due to particle size and distribution (phases or inclusions) leading to the detection of defects or faults with any sufficiently short wavelength.

(3) Loss of energy which occurs in ferromagnetic or ferroelectric materials due to movements of domain walls. In magnetic materials there is a change of elastic modulus with magnetization, micro-eddy currents and a micro-hysteresis effect.

(4) Loss of energy due to movement of interstitial or substitutional atoms. This is the effect described by Snoek (Zener 1948, Snoek 1941).

(5) Effects of dislocations on the damping of vibrations can arise due to comparable diffusion effects requiring the diffusion of atoms between dislocation sites.

(6) Alternatively, ultrasonic waves could be attenuated by processes such as those described by Mason. These involve (i) consideration of low-amplitude vibrational movements of dislocations as similar to the behaviour of stretched strings, (ii) the break away of loops from dislocation atmospheres, (iii) at larger amplitudes the loops can become unstable and generate other loops. This involves anelastic internal friction and plastic flow. Plastic flow can occur by similar mechanisms under direct stress (zero frequency). The dislocation theory can thus be applied to the interpretation of fatigue in general.

(7) Grain boundary, thermally activated, relaxation.

(8) Attenuation due to any other structural elements, e.g. vacancies, ions, molecules or electrons. Thus, high frequency sound waves are attenuated by the conduction electrons at low temperatures.

Attenuation is observed or detected as a continuous process, e.g. specific dislocations are not observed. If,

however, the wavelengths are of the same order as the separation between or the size of the structural units scattering can occur and might lead to loss of energy by diffraction. Alternatively scattering by individual particles might enable them to be recognized provided the right conditions for experimental observation are established.

Electromagnetic and radiation properties

Some electromagnetic behaviour is already covered by the above discussion, but there are other aspects not explicit or implicit in this treatment. The most general principles are the equations which determine the electric or magnetic fields and lead to Maxwell's equations for electromagnetic radiation. These include the principles of mutual and self induction and can be considered without any reference to structure. The magnetic permeability and electric permittivity or dielectric constant, required in practical application are amongst the structural properties already considered.

(1) Electrical resistance can now receive further consideration. The structural effects involved are:

(*a*) Density of 'free' or conducting electrons. The distinction between conductors, semiconductors and insulators must be made here and can be interpreted in terms of band theories, etc.

(*b*) Structure of the solid including anisotropy.

(*c*) General rise of resistance with temperature due to thermal motion.

(*d*) Impurity atoms and any kind of lattice defects or irregularities including grain boundaries can cause scattering.

(*e*) Proceeding downwards in Fig. 3 we expect variations in resistance to be associated with any of these changes or with the presence of gross defects.

In regard to the last of these observations it is realized that in any given test a change in this physical property may indicate a number of possible changes within the material without being able to distinguish between them or locate very accurately. This does not limit the use of resistance methods for general quality control, but limits their use for defect detection.

(2) Proceeding towards higher frequencies, as on the left-hand side of Fig. 1, physical methods become less concerned with fields and currents in the material as conductor (or insulator) and progressively more concerned with the free radiation. All radiation can be attenuated by various processes. It can be absorbed, reflected, refracted, scattered or diffracted according to the conditions. In addition to those effects which depend on crystal properties (as above) there are some other general properties which depend more definitely on the atomic structure and these are as follows:

(*a*) Total absorption or attenuation of radiation on passing through a material is governed by the basic relationship

$$\frac{dI(x)}{I(x)} = n\sigma dx$$

where n is the number of atoms per cubic centimetre and σ is the 'atomic attenuation coefficient'. Considering all the possible processes σ can be the sum of four terms due to (i) photoelectric effect, (ii) scattering, (iii) pair production, (iv) photodisintegration (McMaster 1959, ch. 13). The last two processes only become significant at high energies (gamma and cosmic rays). For x-ray absorption the linear absorption coefficient $\mu (= N\sigma\rho/A$, where $A =$ atomic weight, $\rho =$ density and $N =$ Avogadro's number) is generally employed.

(*b*) The absorption due to the photoelectric effect involves the ejection of electrons from energy states according to the relation

$$h\nu = \Delta E + E_{\text{kin}}$$

where $\Delta E =$ difference in energy level, $E_{\text{kin}} =$ kinetic energy of ejected electron (can give heating effects).

(*c*) The scattering can be without change of wavelength and some of this scattering can give diffraction under suitable conditions or it can be scattered with change in wavelength according to the Compton effect (which becomes less pronounced towards higher energies).

(*d*) The absorption of quanta into molecular vibrational or rotational energy bands is a process based on $h\nu = \Delta E$, but is completely separate from the photoelectric absorption.

(*e*) Any fluorescent process can follow a photoelectric absorption—the emitted frequency is lower than the absorbed frequency due to the kinetic energy fraction.

(3) The effects of the increasing structural complexities as represented in Fig. 3 can be described as

(*a*) Mixing effects, due to the presence of atoms of different kinds. The atomic processes indicated can happen for each atomic species separately. Thus the method of combining absorption coefficients follow.

(*b*) Differential absorption effects. Apart from changes of geometrical cross section the structural effects are (i) different absorption by different phases or constituents on a micro-scale, detectable by micro-radiography, (ii) different absorption in path of a beam due to macro segregation or defects requiring no magnification.

Both processes depend on whether the absorption differences can give measurable differences in contrast and whether an object or area of different absorption properties can be resolved by the detecting or recording device.

A note on the mechanical properties and plastic deformation of metals

From a metallurgical point of view plastic deformation and mechanical properties would require a number of further considerations. The dislocation theories have been particularly successful here (Cottrell 1953, Fisher *et al.* 1956) and progress has been made in connection with the treatment of fracture (Cottrell 1958). Whilst it is generally true that the physical properties considered do not correlate easily with (non-elastic) mechanical properties and the plastic deformation history of a material there are a number of distinct possibilities already apparent. For example, hardness and tensile strength correlate with grain size which correlates with electrical resistance. Plastic deformation affects these properties at least partly through the grain size. Otherwise it increases the number of dislocations and so the ultrasonic attenuation should also be altered.

The effects of large macroscopic faults and defects on mechanical properties is somewhat arbitrary and individual in character, but it is at this level non-destructive tests are most successful. It is the development of smaller submicroscopical defects (micro-cracks) generally which can contribute to an overall level of strength (or weakness) and which individually can become (undetected or unsuspected) sources of failure. The correlation with physical properties here would be most desirable, but is perhaps the most difficult. This field of research needs much more consideration than we are able to give it here.

Conclusion

A summary of this kind is necessarily very general and perhaps over-simplified. It will have served a purpose if it suggests that most physical methods for the examination or testing of materials are interrelated, and secondly, they can be thought of more satisfactorily if the structural aspects (atomic or crystalline) of the material are considered as part of a basic unity of properties and material. Progress in non-destructive testing should certainly follow from a continual appraisal of all of the basic physical properties that are available. We also hope that such progress will be assisted by the proper introduction and use of structural aspects and ideas which serve to explain and interpret the properties.

Acknowledgments

The writer records his indebtedness to a number of colleagues who have from time to time discussed various aspects of this subject with him. Mr. W. C. Heselwood kindly provided a frequency spectrum on which part of Fig. 1 is based. Particular mention should also be made of a very stimulating lecture given in April 1959 in Sheffield by Professor G. A. Homès in which the whole field of physical testing was surveyed.

The author would like to thank Mr. F. H. Saniter, O.B.E., Director of Research, United Steel Companies Limited for permission to publish this paper.

References

BOZORTH, R. M., 1951, *Ferromagnetism* (Princeton: Van Nostrand).

CADY, W. G., 1946, *Piezoelectricity* (New York: McGraw-Hill).

COTTRELL, A. H., 1953, *Dislocations and Plastic Flow in Crystals* (Oxford: Clarendon Press).

—— 1958, *Trans Amer. Inst. Min. (Metall.) Engrs*, **212**, 192.

FISHER, J. C., THOMSON, R., JOHNSON, W. G., and VREELAND, T. (Editors), 1956, *Dislocations and Mechanical Properties of Crystals* (Lake Placid Conference), (New York: Wiley).

GUINIER, A., and FOURNET, G., 1955, *Small Angle Scattering of X-rays* (London: Chapman and Hall).

HINSLEY, J. F., 1959, *Non-Destructive Testing* (London: Macdonald and Evans).

INSTITUTE OF METALS, 1958 a, *Advances in Inspection Techniques as Aids to Process Control in Non-Ferrous Metals Production*, Institute of Metals Special Report No. 24.

—— 1958b, *Vacancies and other Point Defects in Metals and Alloys*, Institute of Metals Monograph and Report Series No. 23.

—— 1959, *J. Inst. Met.*, **87**, 385–458.

LEWIS, D. M., 1951, *Magnetic and Electrical Methods of Non-Destructive Testing* (London: Allen and Unwin).

MCLEAN, D., 1957, *Grain Boundaries in Metals* (Oxford: Clarendon Press).

MCMASTER, R. C. (Ed.), 1959, *Non-Destructive Testing Handbook, Vols. I and II* (New York: The Ronald Press Company).

MASON, W. P., 1958, *Physical Acoustics and the Properties of Solids* (Princeton: Van Nostrand).

NYE, J. F., 1957, *Physical Properties of Crystals* (Oxford University Press).

VAN OUWERKERK, L., 1957, *Non-Destr. Test.*, **15** (5), 298.

RAIT, J. R., and PEISER, H. S., 1957, *Physics of Non-Destructive Testing, Brit. J. Appl. Phys. Suppl. No.* 6, esp. p. S.13.

SNOEK, J. L., 1941, *Physica*, **8** (1), 711.

STANFORD, E. G., and FEARON, J. H., 1958, 1960, *Progress in Non-Destructive Testing, Vols. I and II* (London: Heywood and Company).

WEINBERG, F., 1959, *Progr. Metal Phys.*, **8**, 105.

ZENER, C., 1948, *Elasticity and Anelasticity of Metals* (University of Chicago Press).

The investigation of surface deformation

by L. GRUNBERG, D.Sc., M.Sc.(Birm.), F.R.I.C., F.Inst.Pet., D. SCOTT, F.I.M., and K. H. R. WRIGHT, Ph.D., A.Inst.P., Lubrication and Wear Division, National Engineering Laboratory, Thorntonhall, Glasgow

Paper presented at the Joint Conference of the Non-Destructive Testing Group and the Société Française de Métallurgie, 2nd to 4th May 1960

Abstract

The methods for studying surface deformation used in the National Engineering Laboratory are described. Deformed metal surfaces show an enhanced photoelectric emission which is studied with the aid of open-ended Geiger–Müller counters or electron multipliers. Oxide films stripped from deformed metal surfaces are studied by electron-microscopy and changes in chemical reactivity by radio-active tracers. Surface topography is an important factor in lubrication and wear research and is being studied with the aid of interferometry, phase contrast and electron microscopy. Replica techniques are often applied with advantage.

Introduction

When metallic surfaces come into frictional contact changes occur both in the topography and in the structure of the surface zone subject to deformation. The changes are of great significance for the subsequent behaviour of the contacts and are therefore of considerable importance to research into lubrication and wear. The study of the nature of the surface zone is, however, of wider interest since most industrially produced surfaces are the result of surface deformation, such as cutting, milling, grinding and so on. Although such surfaces soon acquire a surface oxide film, the consequences of the deformation produced in manufacturing are preserved below this film and become of importance when the oxide film is broken. The present paper attempts to outline briefly some of the methods used in the Lubrication and Wear Division of the National Engineering Laboratory for the study of deformed metal surfaces.

Electron emission studies

It has been known for roughly a decade that freshly deformed metal surfaces are capable of emitting electrons in circumstances in which annealed or aged surfaces are completely inert. A review of the subject has recently been published (Grunberg 1958). The original discovery was made by Kramer (1950) who showed that the spontaneous emission of 'exo-electrons' could be observed in the dark. Grunberg and Wright (1953) discovered that, after abrasion, surfaces acquired photoelectric sensitivity at relatively long wavelengths of light.

The electron emission currents obtained from deformed metal surfaces, either spontaneously or through illumination, are rather small, and very sensitive instruments must be used for their detection and measurement. Kramer (1950) used a point counter for this purpose (Fig. 1(*a*)). This consists of a cylindrical cathode, usually kept at earth potential, which contains an insulated plug holding a fine wire, acting as anode and terminating in a fine point or a small metal sphere. High anode voltages up to 4000 v are required, but the instrument can be operated at atmospheric pressure and in air and is therefore readily adapted to various kinds of experiments. Unfortunately point counters are rather unstable and tend to collect dust particles from the atmosphere and this leads to spurious counting. In order to increase stability, open-ended Geiger–Müller counters are used; they can be either of the vertical or of the horizontal type (Fig. 1(*b*), (*c*)). These require a relatively low pressure

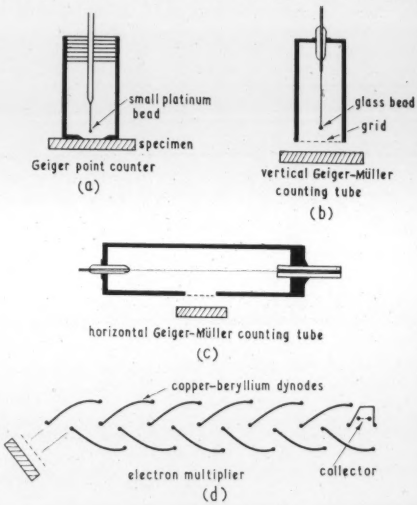


Fig. 1. Types of instruments used for measuring exo-electron emission.

atmosphere (4–10 cm Hg) of argon, hydrogen or helium, and the presence of a 'quenching' gas such as ethyl alcohol or a hydrocarbon gas. Operating voltages up to 1500 v are satisfactory. The specimen should be shielded from the extraneous field effects by a wire grid which is kept at a slight positive potential relative to the specimen and hence draws electrons emitted from the surface into the counter.

The electrons emitted from deformed surfaces have a low energy and are incapable of penetrating even the thinnest counter windows. The specimens have therefore to be kept in the special atmosphere required for the satisfactory operation of the counter. This introduces uncertainties about the effect of the environment on the emission process and recently there has been a tendency to use instruments capable of operating in a high vacuum. Most of these are based on secondary emission electron multipliers (Fig. 1(*d*)). The electrons are accelerated by high tension fields between

dynodes and, since the number of secondary electrons emitted from the dynodes greatly exceeds the number of primaries, high amplification factors can be obtained.

A few examples will illustrate the use to which the measurement of emission currents can be placed. Abraded surfaces have been studied in detail (Grunberg and Wright 1955) and it is of interest to investigate the depth to which abrasion causes deformation in a metal surface. An aluminium specimen was abraded in air; under illumination with light immediately afterwards it gave a high emission current in the near ultra-violet and violet range of wavelengths. After the specimen had been kept for about two days in air the photoelectric sensitivity had completely disappeared, owing to the growth of an oxide film over the surface. Successive layers were etched from the specimen with dilute hydrofluoric acid and the emission characteristics were investigated. Fig. 2 shows the results of this experiment. The photo-

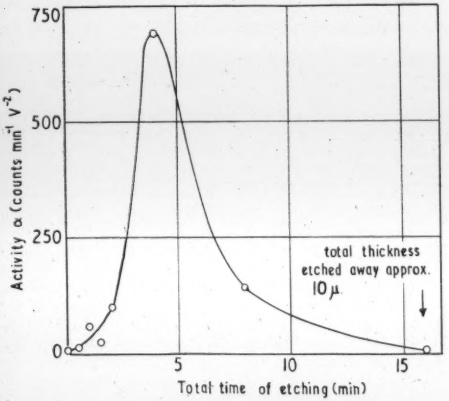


Fig. 2. Emission from an etched aluminium specimen.

sensitivity at first increased and reached a maximum after approximately 2–3 microns had been etched from the surfaces. Beyond this region the activity decreased and reached zero at a depth of about 10 microns. Abrasion, although of considerable practical importance, is not a well-defined deformation process. In order to study the relationship between strain and emission, specimens of aluminium were subjected to tension, under illumination, in a chamber containing a Geiger–Müller counter (Grunberg and Wright 1957). Using annealed specimens covered by an air-grown oxide film of thickness 20 to 50 Å, no photo-sensitivity was observed at strains below 2%. At strains between 2 and 5% the photo-sensitivity appeared and began to increase rapidly with increasing strain (Fig. 3). The increase of photo-sensitivity, as measured by the magnitude of the emission current J, was dependent on the atmosphere in which the experiment was performed. In an atmosphere of argon and ethyl alcohol $J \propto (e - e_0)^3$, and in an atmosphere of air and alcohol $J \propto (e - e_0)^2$, where e is the instantaneous strain and e_0 the initial strain at which the photo-sensitivity appeared. Recent experiments carried out in a high vacuum apparatus (10^{-5} mm Hg) embodying an electron multiplier, showed that the emission tended to 'overshoot' the strain, i.e. the electron current increased after the stressing of the specimen had stopped (Fig. 4).

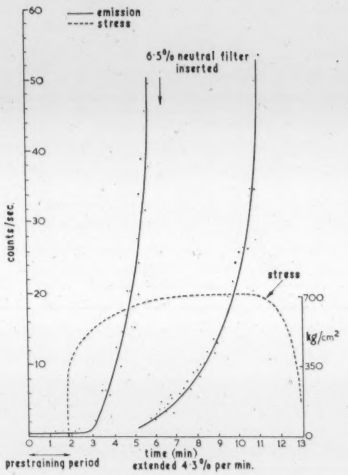


Fig. 3. Emission from extended aluminium.

The interpretation of the results from emission studies is as yet rather difficult, since the phenomenon is not fully understood. It seems likely that vacancies arising in the metal owing to the interaction of dislocations (or at jogs) diffuse to the surface. The high concentration of vacancies in the surface region of the metal and the presence of oxygen may lead to the formation of an oxide film containing oxygen ion vacancies in which electrons are trapped. The high rate of oxidation occurring under these conditions may lead to the thermal emission of some electrons and may thus be the cause of the 'spontaneous' emission reported by several observers. Electrons are held in the traps with relatively low energies and can be extracted by illumination at relatively long wavelengths. Electron emission studies can thus be used for studying both the results of deformation of a metal surface and of the interaction of the latter with oxygen.

Electron microscopy of oxide films

When a metal surface is being deformed the oxide film covering the surface must respond to the deformation in a manner which is compatible with its mechanical properties. Oxide films are rather brittle and the strain imposed by the slip system of the ductile metal substrate could be expected to lead to brittle failure of the film. This phenomenon was studied with the aid of electron microscopy. Annealed aluminium specimens were deformed in tension and then covered with a Formvar film. The composite film of oxide and Formvar was removed from the surfaces by etching with mercuric chloride solution and examined in an E.M.3 (Metropolitan-Vickers) electron microscope. Figs 5 and 6 illustrate some of the results obtained. The slip system in the aluminium had caused brittle failure in the oxide over the slip bands. This causes the exposure to the atmosphere of the highly dislocated metal contained in the slip band (Fig. 5). Occasionally transverse cracks appear (Fig. 6) suggesting that the oxide film has slipped relatively to part of a slip block in the aluminium. Both these observations

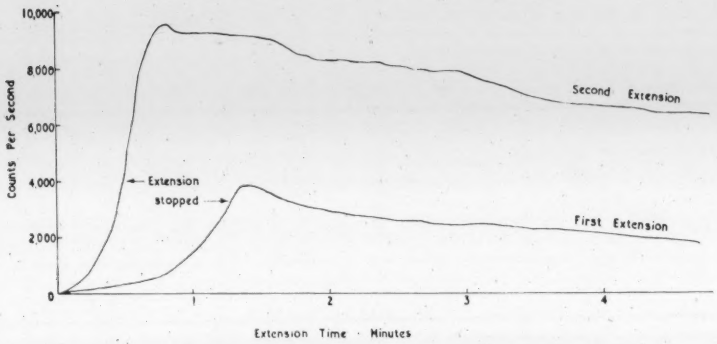


Fig. 4. Variation of emission current with extension time.

lend support to the vacancy mechanism suggested for the emission phenomena. The metal in the slip band may be expected to contain a proportion of defects, including vacancies. The relatively poor adhesion of the oxide film to the metal may be due to an accumulation of vacancies at the oxide–metal boundary.

Reactivity of deformed metal surfaces

The physical changes produced by deformation are likely to change the chemical reactivity of surfaces. This can be seen from the following example. Specimens of aluminium were deformed in tension under solutions of C-14 labelled stearic acid in benzene. The reaction was allowed to proceed for a total period of ten minutes. Fig. 7 shows the amount of stearic acid taken up by the surface as a function of strain and of concentration. Within the short period allowed for the reaction the coverage was relatively small and depended on concentration. The amount of stearic acid taken up increased with strain and this may be due to interaction with strained metal in slip bands exposed by the breaking-up of the oxide film.

Surface topography

In any investigation of surface deformation the original surface has to be studied before changes due to deformation are considered. Several techniques are available for the examination of the topography of surfaces and the method used depends upon the nature of the specimen and the information required from it. Many optical methods allow the specimen to be investigated without destroying it or subjecting it to strain or damage (Scott 1957). Stylus methods provide an immediate numerical characterization of the surface, such as centre line average, but although they give a record of the surface contour quickly and conveniently, they are subject to the uncertainty of surface damage by the stylus, if the surface is soft. The profile microscope which, like stylus methods, gives only a single line profile, may yield insufficient information and it may be necessary to observe many different recordings in order to obtain a comprehensive conception of the surface. Observing an enlarged image of the surface by optical microscopy, although useful, allows no conclusions to be drawn regarding surface roughness, so interferometry is used to reveal and evaluate minute surfaec structures. In the image these appear as deviations in the

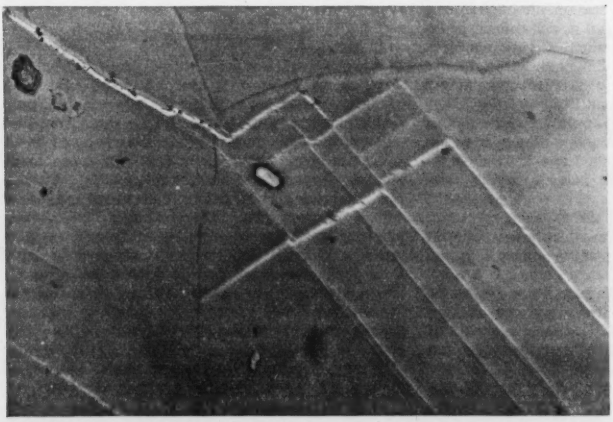


Fig. 5. (×9000) Failure of deformed oxide film.

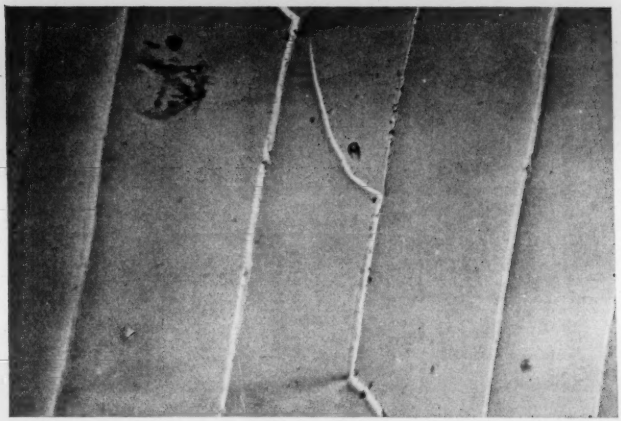


Fig. 6. (×9000) Failure of deformed oxide film.

fringes which represent contour lines connecting all points of the same level. The difference in level between fringes corresponds to half a wavelength of the light used. The magnification is irrelevant and the geometric shape of the

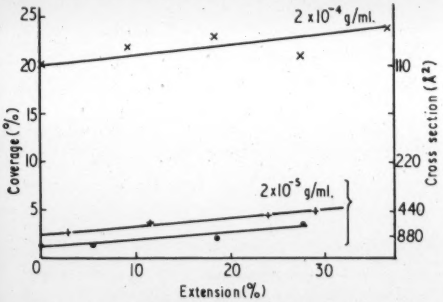


Fig. 7. Adsorption of stearic acid on extended aluminium.
× concentration of stearic acid 2×10^{-4} g/ml.
+, • concentration of stearic acid 2×10^{-5} g/ml.

test piece is not important. The difference between an electrolytically polished and a mechanically polished aluminium surface, prepared to the same average roughness by stylus measurement, which had been used in a friction experiment, is clearly shown by interferometry in Fig. 8 (×50). The electrolytically polished surface, although highly reflecting to light, is undulating, whilst the mechanically polished surface is quite flat. Comparing the deformation or track formed as the result of a single traverse by a hemispherical steel slider on the respective surfaces when oleic acid was used as a lubricant, the greater amount of deformation and build-up is shown on the mechanically polished surface. Definite deformation bands can be seen within the tracks, which suggests that the slider had picked up material which caused deformation of the softer surface.

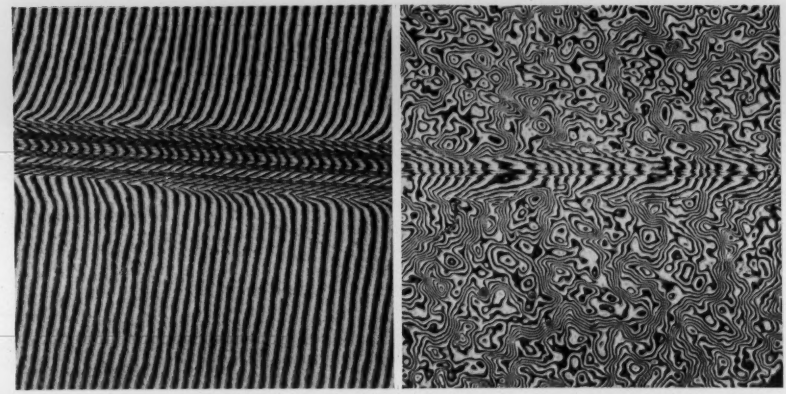
Suitable selection of the method of specimen illumination can extend the versatility of the modern microscope; particular surface features may be emphasized by specific forms of illumination. Some surface irregularities may be more easily observed by dark field or oblique illumination; the irregularities which diffract the light appear light on a dark background. Phase-contrast microscopy may be used to advantage to examine surface deformation as it is very sensitive to changes in surface levels and can reveal fine-scale surface irregularities quite invisible under normal illumination. The true nature of an anodized aluminium surface is revealed by phase-contrast as shown in Fig. 9 (×75). Polarized light may be used to supplement information obtained by phase-contrast and interference techniques.

When a more detailed examination of surfaces is required the electron microscope with its enhanced magnification range may be used (Scott and Scott 1957). Sufficiently small specimens may be examined directly by reflection technique, but usually transmission techniques are used with surface replicas. Fig. 10 (×7500) shows the three-dimensional appearance of a two-stage positive Formvar–carbon replica of a finely ground steel surface. The flowed uneven nature of the surface can be clearly seen. Comparison between a more uniform 'mirror finish' prepared by diamond polishing and by running in, is shown in Figs 11 and 12 (×9000). The electron microscope shows the difference between these surfaces which, by optical examination and stylus assessment, appear to be similar.

The high resolution of the electron microscope is required to study and clarify fundamental phenomena associated with displacement or removal of surface material (Milne, Scott and Macdonald 1957). Electron microscopy is also essential to reveal the initiation of deformation leading to fatigue and creep failure. Fig. 13 (×7500), for instance, shows deformation on the electrolytically polished surface of a copper fatigue specimen, examined after only one tenth of its fatigue life.

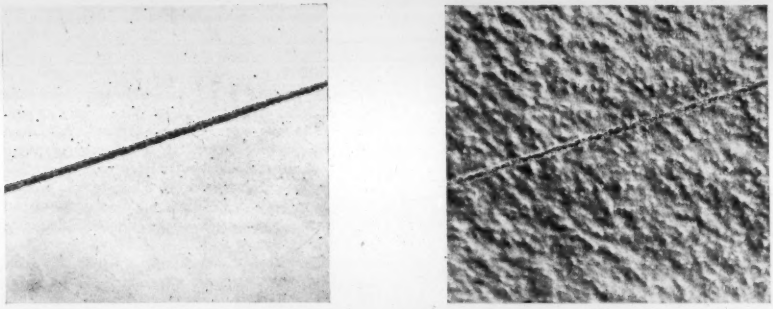
Replica techniques

In order to examine microscopically selected surface areas of large unwieldy test specimens, or to examine periodically the surface of large machine components *in situ*, replica techniques may be used to advantage. Plastic replicas may be made quickly and, to aid microscopic examination and allow investigation by interferometry, a highly reflecting sur-



(*a*) (*b*)

Fig. 8. (×50) Friction trace across a polished aluminium surface: (*a*) mechanically polished; (*b*) electrolytically polished.



(*a*) (*b*)

Fig. 9. (×75) Stylus trace on an anodized surface: (*a*) normal illumination; (*b*) phase contrast.

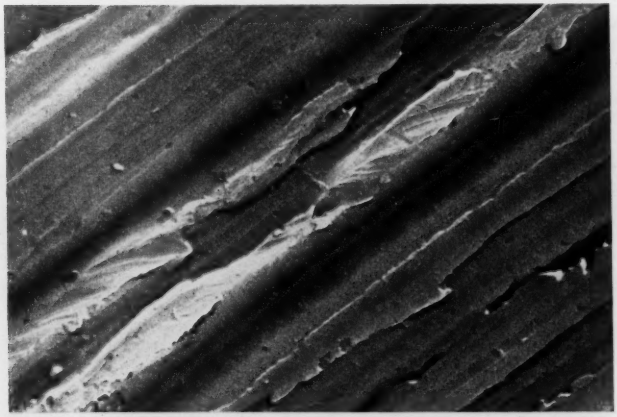


Fig. 10. (×7500) Electron micrograph of finely ground steel surface.

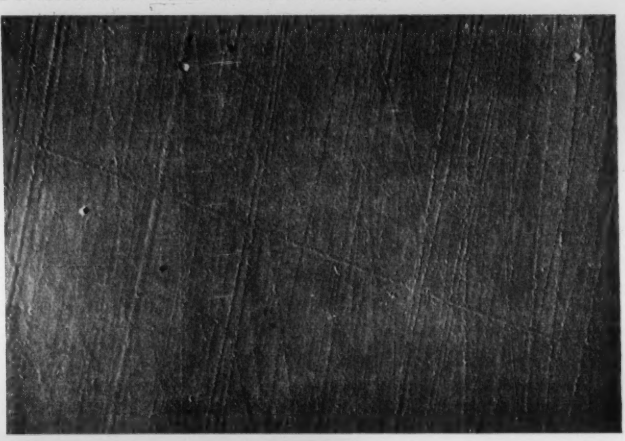


Fig. 11. (×9000) Diamond polished surface.

Fig. 12. (×9000) Run-in steel surface.

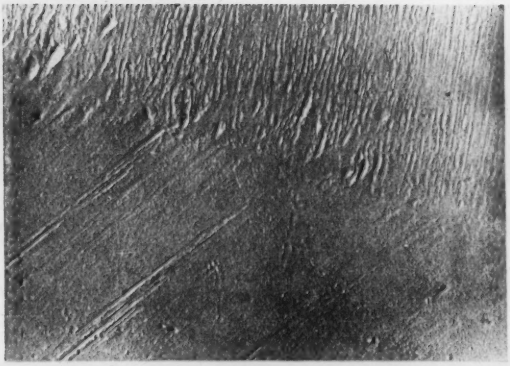
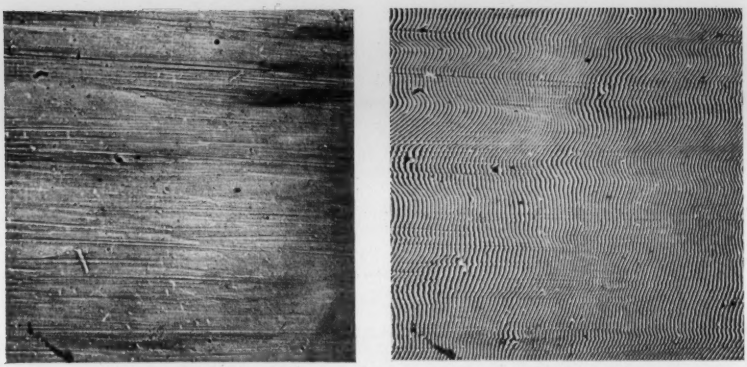


Fig. 13. (×7500) Surface deformation on copper fatigue specimen.



(*a*) (*b*)

Fig. 14. (×38) Replica of worn gear tooth: (*a*) normal illumination; (*b*) interference micrograph.

Fig. 15. (×7500) Fractured steel ball.

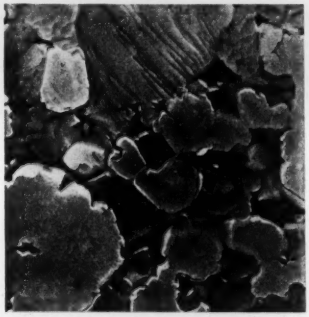


Fig. 16. (×3000) Worn tool tip.

face can be produced by coating the replica *in vacuo* with a thin film of aluminium. Fig. 14 (×38) shows the surface of a tooth of a worn hypoid pinion obtained by replica method. Interferometry shows that the surface is only lightly scored, but is rippled. For the examination of curved surfaces of small radius, pliable replicas made of softened cellulose acetate sheet are useful as they can be flattened out to facilitate focusing of the microscope. These replicas can be examined by transmitted light or by reflected light, with or without a metallic coating to increase reflectivity.

Surfaces which are too rough to be examined satisfactorily by optical microscopy, owing to the limited depth of focus, can be conveniently examined by making use of the great depth of focus of the electron microscope. Such rough surfaces, for example fracture surfaces, may be replicated successfully for electron microscopic examination by means of two-stage Perspex–carbon replicas. The Perspex replica is made thick and strong to avoid strain or tearing when stripping from the rough surface and the carbon replica of the Perspex is made very thin to give maximum contrast and is freed by dissolution of the Perspex. Fig. 15 (×7500) shows a typical electron micrograph of a fractured steel ball, obtained by this technique. Fig. 16 (×3000) shows fine detail, not visible by optical microscopy, on the rough surface of a crater in a cemented carbide tool tip, but in this instance the replica was prepared with Formvar.

Acknowledgment

The work described formed part of the research programme of the National Engineering Laboratory of the Department of Scientific and Industrial Research and is published with the approval of the Director.

References

GRUNBERG, L., 1958, *Brit. J. Appl. Phys.*, **9**, 85.

GRUNBERG, L., and WRIGHT, K. H. R., 1953, *Nature* (*London*) **171**, 890.

—— 1955, *Proc. Roy. Soc. A*, **232**, 423.

—— 1957, *Acta Phys. Austr.*, **10**, 376.

KRAMER, J., 1950, *Der Metallische Zustande* (Göttingen: Vandemhoeck and Ruprecht).

MILNE, A. A., SCOTT, D., and MACDONALD, D., 1957, *Proc. Instn Mech. Engrs, Lond. Conf. Lubrication and Wear*, p. 735, Paper 97.

SCOTT, D., 1957, *Proc. Instn Mech. Engrs, Lond. Conf. Lubrication and Wear*, p. 670, Paper 57.

SCOTT, D., and SCOTT, H. M., 1957, *Proc. Instn Mech. Engrs, Lond. Conf. Lubrication and Wear*, p. 609, Paper 14.

Recent advances in magnetic analysis

by K. HOSELITZ, Ph.D., F.Inst.P., Mullard's Research Laboratories, Salfords, Redhill, Surrey

Paper presented at the Joint Conference of the Non-Destructive Testing Group and the Société Française de Métallurgie, 2nd to 4th May 1960

Abstract

Some recent developments in the examination of materials with the aid of the magnetization curve are described. The well-known method of estimating internal stresses by means of magnetic measurements has recently been much refined so that a quantitatively correct estimate of the internal stresses arising from plastic deformation of nickel and even iron can be made. In special cases even inhomogeneous stresses can be magnetically estimated and the approach to saturation is related to the dislocation density. A fair amount of information about non-magnetic inclusions and precipitates in a magnetic material can be derived from a critical assessment of its coercivity. On the other hand, the measurement of torque curves and of rotational hysteresis allows a comprehensive analysis of the shape, size and density of magnetic precipitates. Finally, the use of magnetic viscosity measurements for structure studies is mentioned.

Introduction

MAGNETIC properties of metals have been used for the examination of structural conditions such as defects, stresses and inclusions in metals and alloys for a considerable time. It is by now well known that the magnetic properties are sensitive functions both of the composition and of the structure of a metal or alloy. In many investigations magnetic measurements have been used to supplement information gained by other methods and this is thought to be the best method of using them. It is also well known that there are structure sensitive and structure insensitive magnetic properties, the main latter ones being the saturation magnetization and the Curie temperature. A fair amount of research has taken place to describe the methods which can be used to determine composition and phase constitution of alloys by means of measurement of primary magnetic properties; this has usually been called thermomagnetic analysis. This review will not deal with this subject, (*a*) because it is preferable to enter into some detail in a small field rather than treat a large number of things superficially, and (*b*) because thermomagnetic analysis has been recently described elsewhere (e.g. Hoselitz 1960) and no really major advances have taken place since these articles were published.

However, the theory of the magnetization curve has developed very substantially during the last ten years and this development has enabled a much greater insight to be obtained into the dependence of magnetic properties on structure. Moreover, these advances have been such that the examination of magnetic metals by means of the magnetization curve can now yield some unique information and, although the interpretation of measurements is difficult, this information can be very precise indeed.

The magnetization curve

The general features of the magnetization curve are well known. Attempts at summarizing the way in which the magnetization curve can be used for structure analysis have been made (e.g. Libsch and Conard 1954) and therefore it is not necessary to go into too great detail. Briefly, then, a magnetic material in the demagnetized state consists of a number of domains which are magnetized in different directions so that the vector sum of the magnetization is zero. Application of a field can do one or both of two things. It can move domain walls in a direction so that more favourably magnetized domains grow in volume at the expense of less favourably magnetized ones, or it can rotate the spin direction, i.e. the direction of magnetization inside the individual domains, to coincide more nearly with that of the applied field. It has been shown by earlier workers that the domain walls can have stable positions from which it is difficult to move them and to which they will tend to return provided they have not moved too far away (Becker and Döring 1939, Néel 1946). The strength with which domain walls are kept in such stable places is related either to the stress pattern in the magnetic material or to the pattern of non-magnetic or other magnetic inclusions. Thus, the initial permeability, for instance, which relates to magnetization processes of a reversible nature in very small fields is often a measure of the stability of such domain walls and can give an insight into the stresses or inclusion pattern in the material. As magnetization proceeds, we usually come to a region where domain wall movements or rotations take place irreversibly and where, upon reduction of the field, the magnetization does not return to its original value. The mean field for such irreversible changes is the coercive force. This also is related to the structural pattern in the material and measurements of the coercive force can be used to study imperfections.

Finally, it is important to realize that the approach to saturation in high fields is mainly due to rotations after all domain walls have moved out of the material. These rotations take place against anisotropy forces inside the material and, therefore, the ease or difficulty with which a material can be saturated, gives a clear measure of these anisotropy forces.

The most important contribution to the theory of the magnetization curve in the last few years has been the recognition that magnetic particles can be so small that the formation of a domain wall is energetically less favourable than the maintenance of the external magnetostatic field, which belongs to a uniformly magnetized element (Stoner and Wohlfarth 1948). In such particles, the magnetization can only change by rotation and this rotation can be irreversible if there are certain preferred directions which can arise from crystalline symmetry, from the shape of the particles or from any external or internal directional stresses. Moreover, it has been shown that in such particles the rotation

of magnetization may be completely coherent, that is, all the spins remain parallel throughout the particle during changes in magnetization and they just rotate in unison, or that spins in one part of the particle can rotate before those in another part have completely rotated by the same amount or in the same direction (Jacobs and Bean 1955).

The most general case is that in which the spins rotate independently. However, this is opposed by exchange forces which tend to keep all spins parallel and hence the solution of the problem of the actual way in which the magnetization changes in such a material presents a complicated boundary value problem which can only be solved in very special cases. Solutions for such cases are possible for cylinders, plates, spheres, ellipsoids of revolution and simple assemblies or combinations of these (Brown 1959, Aharoni 1959).

The magnetization curve of a heterogeneous material can always, in principle, be interpreted as a combination of domain wall movements and rotations and we now describe, with a few examples, how the precise theory of the magnetization curve can be used in some cases to interpret the structural state of some magnetic metals.

Some examples of modern analysis of the materials using the magnetization curve

One of the oldest relations in the theory of the magnetization curve has been that which has related internal stresses to either the initial permeability or the coercive force or the approach to saturation (Becker and Döring 1939). In some cases, magnetic analysis has been used to attempt to determine internal stresses. Qualitatively good agreement is obtained, especially for materials in which the magnetostriction constant is large and in which, therefore, the anisotropy energy which arises from stresses can be large compared with other contributions to the magnetization energy. However, if this agreement is examined more quantitatively it is found to be out by a factor of 3 or more even in a most obvious case, such as, for instance, pure nickel. This is not due entirely to the lack of success of the theory of the magnetization curve, but partly to the fact that internal stresses had not been critically analysed by metal physicists. The situation has changed substantially in the last few years mainly because Taylor (1938), and especially Greenough (1949), have analysed homogeneous internal stresses in polycrystalline materials in terms of directions of the individual grains. Briefly, this analysis shows that if a polycrystalline rod is homogeneously stressed, the distribution of internal stresses in the individual grains depends on their orientation with respect to the stress direction. In fact, if the lattice spacing, which is a measure of the internal strains, is measured by x-rays and that part which is due to homogeneous stresses is plotted against the angle which the grains present to the x-rays, this distribution is very close to that which would be theoretically expected on the basis of the Taylor and Greenough theory. It is therefore possible to take the line broadening of a polycrystalline material which is a measure of the spread of lattice spacings and determine the mean value of internal stresses much more accurately now than was possible previously. It is further realized that since the deformation of individual grains depends on the angle which they make with the stress direction, so will the change in magnetization, if it is due to magnetostriction.

In pure nickel the case is simple because the magnetostriction is almost entirely isotropic and, therefore, for any given stress in any direction the change in magnetization is directly proportional to the applied stress. In iron, however, the magnetostriction is very anisotropic, so much so that it is positive in the (111) direction and negative in the (100) direction.

However, an analysis of the changes in magnetization as a function of orientation in a homogeneously stressed polycrystalline bar of iron can be made by considering the magnetostriction as well as the resultant stress for any given grain as a function of direction, and it has lately been possible, in both nickel and iron, to relate the internal stresses arising from homogeneous plastic deformation to the magnetic properties (Reimer 1956). Fig. 1 shows the result for

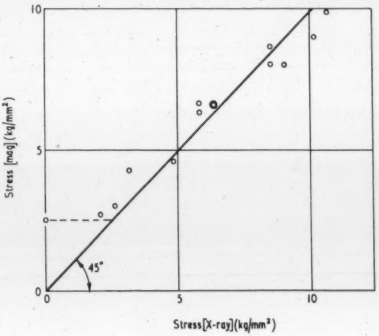


Fig. 1. Comparison of internal stress determined from x-rays with that derived from measurements of magnetization energy in nickel.

a bar of nickel. The stress derived from x-ray line broadening using the Greenough–Taylor theory is plotted horizontally and the stress derived from magnetization curve measurements (in this case, the energy of magnetization) is plotted vertically.

The point for zero x-ray stresses indicates that here no correction has been made for the crystalline anisotropy. ($K_1 = 5 \times 10^4$ corresponds to approximately 10 kg mm^{-2} in pure nickel.) Otherwise, as can be seen, the two methods given show surprisingly corresponding results. In iron the crystalline anisotropy is very much larger than the stress anisotropy energy (corresponding to $\sim$80 kg mm^{-2}) and, therefore, the magnetic data have been obtained by measuring the difference between the magnetization curves of the unstressed and the stressed material. Thus the crystalline anisotropy is directly accounted for and the measurements given only apply to that contribution which is due to the stresses. The experimental results in the case of an iron specimen show very fair agreement up to fairly high internal stresses (Fig. 2). In both cases, the ideal state is indicated by the solid line at 45° to the coordinate axes.

The results of these researches show clearly that if taken seriously and if examined sufficiently critically, magnetic stress analysis is possible and can lead to very accurate determination of homogeneous internal stresses. The distribution of stresses must, no doubt, have an influence on the magnetic values and it must be concluded from examination of the theory that stresses of the third kind, which are inhomogeneous grain boundary stresses and dislocation stresses, are much less readily analysed by magnetic methods.

However, this has been attempted by a number of workers recently and some preliminary success has been achieved. It would lead a little too far to go into the entire theoretical

background, especially since work is still proceeding and the whole matter is, at the moment, in a state of sensitive development. It is, however, of value to indicate that the approach to saturation in high fields when no domain walls are present in the material must be a measure of the difficulty with which

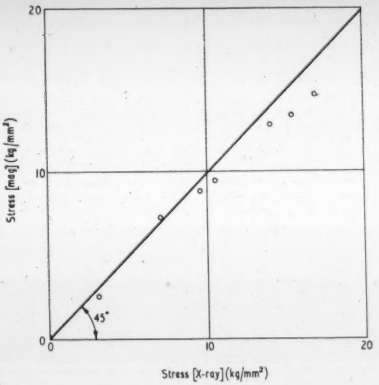


Fig. 2. Comparison of internal stress determined from x-rays with that derived from measurements of the magnetization energy in iron.

the magnetization can be rotated into the field direction. Near disturbance centres, e.g. dislocations, etc., there are very high amplitude short wavelength stresses, which have the effect of making the approach to saturation extremely difficult in small elements of volume. If we have no grain boundaries, we can attribute all these to strains arising from dislocations and a preliminary analysis of the approach to saturation with these large localized strains can be made. In general, the theory of the magnetization curve indicates that the slope of the magnetization curve in high fields should be proportional to $1/H^3$. With large amplitude stresses in local areas a term with C_2/H^2 can arise and C_2 should depend on the dislocation density and be proportional to the plastic stress, i.e. according to dislocation theory it varies with the square root of the dislocation density. Experiments on single crystals of nickel have been carried out by Dietrich and Kneller (1956). Fig. 3 shows two of their curves. From these curves it appears that with crystals which were stressed along one of the principal axes or in a direction near to the

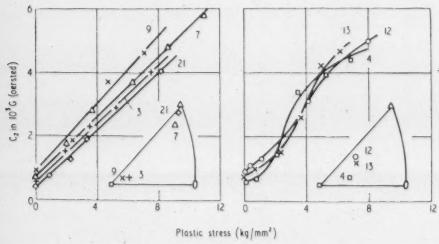


Fig. 3. Variation of the $1/H^2$ term (C_2) with plastic stress in single crystals of nickel. (The numbers on the curves refer to different crystals whose orientation is shown in the insets.)

direction of the principal axes, this theory gives very good agreement because the constant C_2 is very nearly a linear function of the plastic stress. When the stress is applied in intermediate directions, linearity is not so good but can still be considered to be quite fair. This work is continuing and some very promising further results are to be expected from a comparison of dislocation conditions with magnetic properties (Seeger and Kronmüller 1960).

Now let us consider inclusions or precipitates. The magnetization curve theory has already, in its early days, related coercive force to inclusions. It was pointed out that inclusions would be acting on domain wall movement via resulting stresses. Kersten (1943) found that an important contribution to the domain wall energy could arise if the domain walls were considered to be attached to inclusions because of an effect similar to that of surface tension which makes soap bubbles stick to a grid or net. Néel (1946) reconsidered the Kersten theory and found that it was wrong. He found that the coercive force was not due to surface tension effects, but to magnetostatic fields which can arise at the surface of the inclusions. This whole question has been taken up in a very rational way by Dijkstra and Wert (1950) who have made plausible calculations of these effects in simple cases and have given a resulting theory of the coercive force due to domain wall movements in a magnetic medium containing non-magnetic inclusions. The calculation is relatively easy for spherical particles and has been given for such by Dijkstra and Wert. The resulting coercive force depends on whether the domain wall thickness δ is greater or smaller than the diameter of the included spheres d or whether the two are approximately equal. The theoretical curves arrived at by Dijkstra and Wert for the domain wall coercive force of iron with spherical inclusions are shown in Fig. 4, together

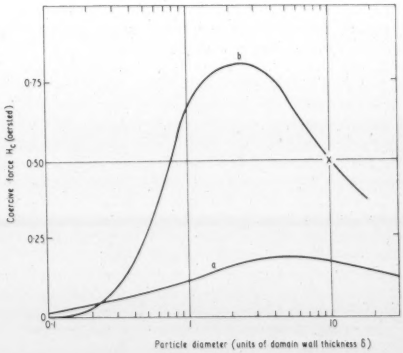


Fig. 4. Theoretical coercive force in iron with spherical non-magnetic inclusions: (*a*) Kersten soap bubble model, (*b*) Dijkstra and Wert theory.

with the coercive force as calculated by the earlier Kersten soap bubble model. Here one should note that the maximum coercive force is reached for a particle diameter d of about 2δ (where δ is the domain wall thickness) and amounts to approximately 1 Oe. Experiments on iron containing carbon in solution were made by these authors and the coercive force plotted as a function of the particle diameter for spherical cementite precipitate in various concentrations and sizes. The result for a volume fraction α of 0·003 of non-magnetic

inclusions is shown in Fig. 5, where it should be noted that the measured maximum is at $d \simeq 1200$ Å. $H_{c(max)}$ has a magnitude of about 2 Oe. According to a theory of diffusion worked out by Zener and checked by electron microscope observations at Purdue University (see Dijkstra and Wert

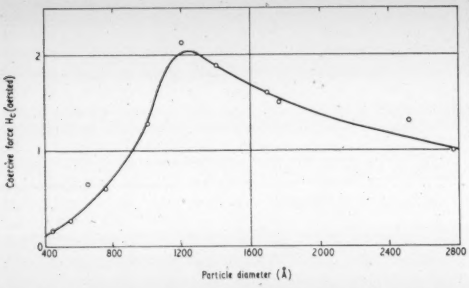


Fig. 5. Experimental variation of coercive force in iron with spherical non-magnetic inclusions.

1950) the particles are spherical and their number can be estimated from the carbon concentration and annealing time. Magnetic theory shows that in iron the domain walls are about 1000 Å thick and, hence, the agreement between theory and experiment is good to within a factor of 1 to 2. This is considerably closer than has ever been observed before and the Dijkstra and Wert results are the first confirmation of the inclusion theory accompanied by direct observation of the inclusions.

Exciting results of a precipitation experiment can be derived by a consideration of the Stoner and Wohlfarth (1948) fine particle theory of single domain magnetization. Many examples of the validity of this theory can be quoted. Some work carried out by Berkowitz and Flanders (1957) at the Philadelphia Franklin Institute has here been chosen because these authors have exploited the theory in a rather profound and illustrative way.

Berkowitz and Flanders examined a nickel–gold alloy of about 25% nickel, quenched from a high temperature and annealed at 400° C. It was known from previous experiments that in such an alloy a ferromagnetic precipitate would come down in a non-magnetic matrix. Berkowitz and Flanders examined this precipitate in its initial stages, i.e. after about 1% of the equilibrium amount had separated out. They were working with single crystals and measured torque curves after this initial precipitation treatment. The method of torque curve analysis evaluated from the detailed tables given by Stoner and Wohlfarth can give information about the precipitate as had been done by Bean and Meiklejohn (1956). It is clear that in an assembly of isolated single domains which are all orientated with easy axes in a given direction, the magnetization will exert a torque on this assembly in high fields. If, however, the precipitate particles are not aligned, no torque will be seen in high fields.

In low fields, however, a torque will be measured which will depend on the direction of rotation of the sample during the recording of the torque curve as a function of angle because each particle tends to remain magnetized in its easy direction until a critical angle is exceeded at which it reverses suddenly. Therefore, the torque at low fields of such random particle assemblies tends to be constant and negative. At intermediate fields an angular distribution of torque may be observed, but the distribution is asymmetric with regard to the torque axis and at very high fields the only distribution remaining will be that due to any particle alignment, and any torque hysteresis will have disappeared. The rotational hysteresis loss, i.e. the difference in torque curves in the two directions of rotation, has been calculated by Bean and Meiklejohn, and their theoretical curves are shown in Fig. 6, both for assemblies of particles with coherent magnetization rotation and with rotation that follows one particular non-coherent mode called the 'fanning' mode. The figure shows that the analysis allows both the determination of the anisotropy field, i.e. the shape or crystal anisotropy of the particles, as well as conclusions about the state of alignment.

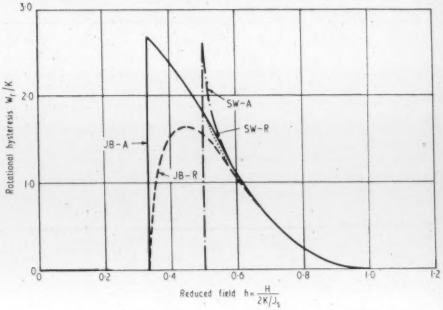


Fig. 6. Rotational hysteresis loss W_r as a function of field in units of 'anisotropy field' $2K/J_s$.

SW, Stoner and Wohlfarth coherent rotation; JB, fanning mode of rotation; A, aligned particles; R, random particles.

The analysis by rotational hysteresis has been carried out by Berkowitz and Flanders on the precipitate magnetic particles in a non-magnetic matrix and some of their experimental results are shown in Fig. 7 for two annealing times, i.e. for two different particle distributions. The torque curves were taken both at room temperature and at liquid nitrogen temperature. The field dependence of the rotational hysteresis is shown in the figure, and can be compared with the

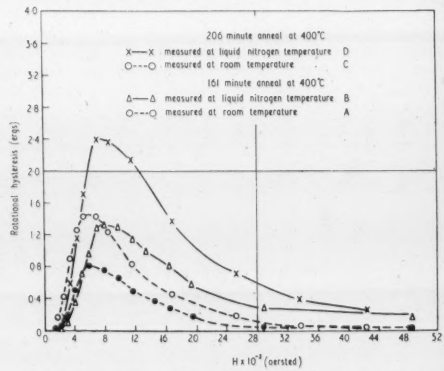


Fig. 7. Experimental rotational hysteresis curves for four specimens of magnetic precipitate particles (nickel) in a non-magnetic matrix (gold).

theoretical curve of Fig. 6. Whilst the experimental curves do not coincide with the theoretical ones they show a strong resemblance and Berkowitz and Flanders were able to analyse the curves in terms of a distribution of particles with different properties. Fig. 8 shows the result of their analysis in terms

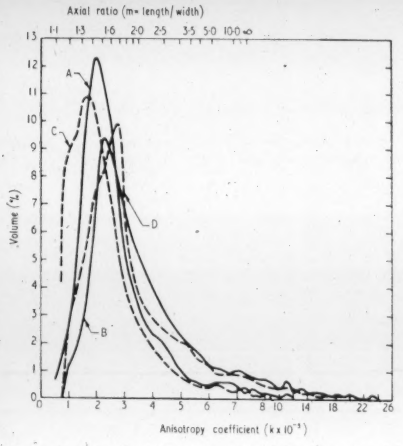


Fig. 8. Rotational hysteresis results of Fig. 7 analysed in terms of particle anisotropy, i.e. axial ratio.

of distribution of anisotropies in the various particles at both temperatures, demonstrating a satisfactory degree of agreement between the results of the analysis for the two tempering times. The analysis shows that the majority of particles have anisotropy coefficients of the order of 2×10^{-5} which would correspond to a value for the precipitate particles of about 1·5 for the ratio of lengths of the long to the short axis. There was corroborative evidence in the experiments that this analysis corresponds with the facts. Examination of these results enabled an insight to be obtained into the way in which the particles grow during precipitation, and it has been concluded that at a lower temperature the crystalline anisotropy of the particles makes a significant contribution. This fact can be seen from Fig. 8.

From the torque curves measured at various fields one can derive the rotational hysteresis integral

$$R = \int_0^\infty \frac{W_r(H)}{I_s} d\left(\frac{1}{H}\right)$$

and the value of this allows a conclusion as to whether the magnetization is coherent or not. The rotational hysteresis integral has values which vary from 0·4 to 10 for different mechanisms depending on the shape of the particles and on whether the rotation is coherent or not. An example of the values obtained for two mechanisms is shown in the Table. Experience in the comparison of theory and experimental results of particle magnetism soon enables us to assess what particle shape we are dealing with when we know the hysteresis integral.

All the magnetic evidence in this case was used to give the following consistent but not necessarily unique picture of the precipitation process. The precipitated particles fall into two groups. One group, the larger one, of nearly spherical

Rotational hysteresis integral. $\int (W_r/I_s) d(1/H)$.

Theoretical values characteristic of mechanism of magnetization and state of alignment

Model	*Alignment*	*Value*
Coherent Rotation (C) or (B)	Random (Sph.)	0·380 ± 0·005
	Aligned or Random (Pl.)	0·415 ± 0·005
Chain-of-Spheres with Fanning (A)	Random (Sph.)	1·02 ± 0·01
	Aligned or Random (Pl.)	1·54 ± 0·01

shape, size 600–2000 Å, axial ratio $m < 1·5$; in these the magnetization rotates incoherently and their speed of growth is fairly rapid. The other smaller group consists of acicular particles, with axial ratio greater than 2, diameter of 850 Å for $m = 2$ to 2500 Å for $m = \infty$, growing more slowly and having a coherent rotational mechanism. A portion of these is oriented along a [001] axis, the remainder are all oriented at random.

Finally, another type of magnetic measurement should be indicated very briefly which can give information about particles or heterogeneous aggregates in alloys. This is the measurement of magnetic viscosity. If one measures the magnetization in a magnetic material and then applies a small change in field, there may be a time dependent variation in magnetic intensity. In general it is found that the magnetic intensity varies logarithmically and the constant of magnetic creep S, in the equation $I = S \log t + C$, can give information about the distribution and size of any domains with unstable magnetization. The energy for reversal in any given domain is related to the coercivity and to the internal field to which the particle is exposed, i.e. that arising from surrounding domains. Spontaneous reversal can take place if the thermal energy kT is approximately equal to this energy of reversal. Under certain special conditions the reversal energy of a particle can be such that it is equal to the thermal energy at some practically attainable temperature, and in such cases the magnetic viscosity can give some structural information (Biedermann and Kneller 1956). An analysis of this sort has been carried out in one or two special cases, especially in the case of the controversial copper–nickel–iron alloys and it has been found that in this case in a certain metallurgical state, the particle shape, size and composition could be determined from magnetic viscosity data. The criterion of the applicability of this method, once the particle character of the precipitate has been established, lies in the temperature variation of the coercive force as compared with that of the magnetic intensity.

Full details here would lead a little too far into the theory but the matter is well understood and the above remarks are intended to indicate that sometimes quite unusual properties, such as the temperature dependence of the coercive force, may have to be measured in order to arrive at data permitting precise magnetic analysis.

There are more ways in which magnetic measurements can help to identify inclusions, precipitates or stresses. In this review only a few isolated aspects of modern structure analysis using magnetic methods have been described. However, it has been shown that the state of the theoretical interpretation of experimental magnetic measurements has advanced very greatly in the last ten years and enables us now to give some

very precise information about structural properties in many cases.

Acknowledgments

I thank the Directors of Mullard Limited for permission to published this paper.

References

AHARONI, A., 1959, *J. Appl. Phys.*, **30**, 705.

BEAN, C. P., and MEIKLEJOHN, W. H., 1956 (unpublished, but cited by JACOBS, I. S., and LUBORSKY, F. E., 1957 in *Conference on Magnetism and Magnetic Materials*, American I.E.E., New York, p. 145).

BECKER, R., and DÖRING, W., 1939, *Ferromagnetismus* (Berlin: Springer).

BERKOWITZ, A., and FLANDERS, P. J., 1957, *Franklin Institute Report No. F*-2482.

BIEDERMANN, E., and KNELLER, E., 1956, *Z. Metallkde*, **47**, 289, 760.

BROWN, W. F., 1959, *J. Appl. Phys.*, **30**, 625.

DIETRICH, H., and KNELLER, E., 1956, *Z. Metallkde*, **47**, 672, 716.

DIJKSTRA, L. J., and WERT, C., 1950, *Phys. Rev.*, **79**, 979.

GREENOUGH, G. B., 1949, *Proc. Roy. Soc. A*, **197**, 556.

HOSELITZ, K., 1960, *The Physical Examination of Metals*, Ed. B. Chalmers and A. G. Quarrell, 2nd edition (London: Edward Arnold), p. 225.

JACOBS, I. S., and BEAN, C. P., 1955, *Phys. Rev.*, **100**, 1060.

KERSTEN, M., 1943, *Grundlagen einer Theorie der ferromagnetischen Hysterese und der Koerzitiv Kraft* (Leipzig: Hirzel).

LIBSCH, J. F., and CONARD, G. P., 1954, *Relations of Properties to Microstructure*, A.S.M. Cleveland, p. 233.

NÉEL, L., 1946, *Ann. Univ. Grenoble*, **22**, 299.

REIMER, L., 1956, *Beiträge zur Theorie des Ferromagnetismus und der Magnetisierungskurve*, Ed. W. Köster (Berlin: Springer), p. 141.

SEEGER, A., and KRONMÜLLER, J., 1960, *J. Phys. Chem. Solids*, **12**, 298.

STONER, E., and WOHLFARTH, E. P., 1948, *Phil. Trans. Roy. Soc.*, **240**, 599.

TAYLOR, G. I., 1938, *J. Inst. Metals*, **62** 307.

Frequency dependence of induced currents

by P. GRANEAU, B.Sc.(Eng.), A.M.I.E.E., Research Laboratories, The British Insulated Callender's Cables Limited, 38 Wood Lane, London, W.12

Paper presented at the Joint Conference of the Non-Destructive Testing Group and the Société Française de Métallurgie, 2nd to 4th May 1960

Abstract

Attempts to solve Maxwell's equations for regions inside electrical conductors generally meet with insuperable mathematical obstacles, unless a uniform magnetic field can be assumed to exist. In many practical instances the field will be non-uniform, since it is generated by currents flowing along wires. The lack of formulae for realistic geometrical arrangements is keenly felt in two cases. They are the prediction of eddy-current power losses and the development of electromagnetic methods for non-destructive testing.

A metallic body can be considered to consist of an infinite number of closed filamentary circuits, coinciding with the streamlines of induced currents. It is then shown, for the most general case, how to express the current at any point in the metal by an infinite series of increasing powers of the frequency of a sinusoidal current flowing in an energizing filament. The coefficients of the series are functions of conductor geometry, electrical conductivity and magnetic permeability, but not of the current itself. This opens the way to measurements on scale models, the long numerical calculation of coefficients by digital computers and the tabulation of normalized coefficients for geometrically similar systems.

Some possible causes of the 'anomalous eddy-current loss' in magnetic laminations are briefly reviewed in order to illustrate the differing opinions that are still being held on this subject. It would, therefore, seem advisable to re-examine the methods of measuring eddy-current losses with reference to the theory of coupled circuits, and thereby determine whether the classical formulae are sufficiently good approximations. Experiments for this purpose are discussed.

Rubber models of yarns and cords; the 'doubling' of single rods

by N. WILSON,* B.Sc., and L. R. G. TRELOAR,* Ph.D., D.Sc., F.Inst.P., The British Rayon Research Association, Wythenshawe, Manchester 22

MS. received 28*th September* 1960, *in final form* 15*th November* 1960

Abstract

The paper is concerned with the calculation and measurement of the stresses involved in the twisting together of two cylindrical rubber rods, which may be taken as a model of a two-ply cord. Two cases were examined: in the first the rods are initially without twist, while in the second they contain initial twist. The forces measured are the couple about the 'cord' axis and the tension along this axis, as functions of 'cord' twist and 'cord' axial extension. The forces are calculated on the basis of Rivlin's equations for the stresses in a rubber cylinder subjected to combined axial extension and torsion, corrections being introduced for the effects of lateral pressure between 'plies' on the tensile force, the non-circularity of section resulting from this lateral pressure on the couple due to torsion and bending, and for the effect of torsion on the bending couple.

For the case of zero initial twist in the rods the agreement between calculated and observed forces is good. When initial twist is present, however, there are significant differences, which are not fully understood.

1. Introduction

In any theory of the physical properties of yarns and cords certain basic assumptions must be introduced regarding (*a*) the geometrical disposition of the filaments or other elements of the structure, (*b*) the physical properties of the filament material, and (*c*) the specification of the stresses and strains in the structure. On account of the complexity of the problem, the verification of a theory by experiments on actual textile yarns is not easy, firstly because their geometrical structure is fundamentally irregular (Riding 1959), and secondly because the elastic properties of typical textile fibres are very imperfect, the stresses not being uniquely determined by the strains.

In order to examine the basic elements of the theory of the stresses in twisted filament assemblies, it has been thought desirable to examine certain simpler systems in which these difficulties are eliminated. For this purpose experiments have been carried out with model yarns, employing vulcanized rubber filaments. This material has the advantage of possessing almost perfect elastic properties, which can be simply specified. In addition, its low modulus and high extensibility permit the use of large 'filaments' which can more easily be disposed in a geometrically definable manner.

The present paper is concerned with the measurement of the stresses in a system comprising two cylindrical rubber rods twisted together, as in the formation of a two-ply cord. Two cases are considered: in the first the two rods are initially free from torsion, while in the second the rods contain initial twist (analogous to singles twist in two-ply cord), before being twisted together. For each of these systems the stresses measured are the tension along the 'cord' axis and the couple about this axis, for various values of axial extension and 'cord' twist. These are compared with corresponding values calculated theoretically, using the measured values of the parameters defining the geometry of the system and of the elastic constants of the rubber.

* Now at the Shirley Institute, Didsbury, Manchester 20

2. Theoretical relations

2.1. *Torque-torsion relation for single rod.* For a cylinder of rubber of unstrained radius b_0 extended to λ times its original length and subjected to a torsion τ about its axis (measured in radians per unit strained length), the total axial force F_1 and axial torque M_1 may be satisfactorily represented by the equations (Rivlin and Saunders 1951, Treloar 1958).

$$F_1 = 2\pi b_0^2\left(\lambda - \frac{1}{\lambda^2}\right)\left(C_1 + \frac{C_2}{\lambda}\right) - \tfrac{1}{2}\pi\tau^2 b_0^4\left(C_1 + \frac{2C_2}{\lambda}\right) \quad (1)$$

$$M_1 = \pi\tau b_0^4\left(C_1 + \frac{C_2}{\lambda}\right). \quad (2)$$

In these equations the constants C_1 and C_2 define the elastic properties of the rubber in a large deformation of the most general type. (Individually these constants have no analogue in the classical, i.e. small-deformation, theory of elasticity, but the quantity $2(C_1 + C_2)$ is equivalent to the modulus of rigidity in simple shear.)

2.2. *The geometrical parameters.* By analogy with the nomenclature of textile cords, the assembly of twisted rods will be referred to as a 'cord', and the individual component rods as 'plies' (Fig. 1). The application of Eqns (1) and (2)

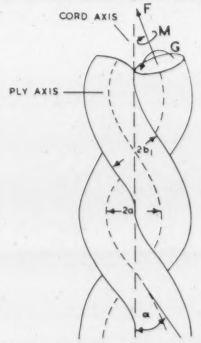


Fig. 1. Force diagram of corded rods.
F, tensile force on rod axis; *M*, couple due to torsion; *G*, couple due to bending.

requires a knowledge of the ply axis length λ (referred to unit unstrained length) and of the torsion τ in the ply. These are obtainable from the geometrical parameters defining the twisted state and the initial torsion in the single rods.

The value of λ is given by

$$\lambda = \lambda_c \sec \alpha \tag{3}$$

where λ_c is the ratio of the cord axis length to the length of the unstrained single rod, and α is the cord helix angle. The latter is related to the cord twist N (turns per unit length of cord axis) and cord helix radius a by the equation

$$\tan \alpha = 2\pi Na. \tag{4}$$

Since λ_c and N are directly measurable, while a is assumed equal to one-quarter of the outside cord diameter (Fig. 1), the ply axis extension ratio λ is thus determined.

The absolute torsion τ in the ply is given by (Treloar 1956)

$$\tau = \frac{\tau_0}{\lambda} + \frac{1}{a} \sin \alpha \cos \alpha \tag{5}$$

where τ_0 is the single rod torsion (in radians per unit unstrained length). The first term on the right-hand side of Eqn (5) represents the contribution due to the initial singles torsion, while the second represents the contribution due to the tortuosity of the ply axis.

2.3. *Components of force and couple in ply.* Consider any plane section of the ply normal to the ply axis. The external forces acting on this plane to maintain the state of equilibrium are (i) a tensile force F normal to the section, (ii) a couple M about the ply axis and (iii) a couple G (bending moment) about an axis normal to the plane of curvature. These forces are shown in Fig. 1, their positive senses being indicated by the arrows.

Eqns (3), (4) and (5) define the values of λ and τ, the extension ratio and twist in the ply, in terms of the measured quantities λ_c, N and a. To obtain the force F_1 acting across the section of the ply in the direction of the ply axis, and the couple M_1 about the ply axis, these values are introduced into Eqns (1) and (2). This assumes that these equations remain valid when the axis of the rod is no longer straight, an assumption which is strictly justified only when the curvature is small.

The tensile force F_1 assumed to act along the ply axis has components $F_1 \cos \alpha$ in the direction of the cord axis and $F_1 \sin \alpha$ in the direction perpendicular to the cord axis and to the radius to the ply axis helix (circumferential component). The total force F_c along the cord axis due to the two plies is thus

$$F_c = 2F_1 \cos \alpha. \tag{6}$$

The circumferential component $F_1 \sin \alpha$ gives rise to a total couple about the cord axis of amount

$$2F_1 a \sin \alpha. \tag{7}$$

The couple M_1 about the ply axis may be resolved into components $M_1 \cos \alpha$ about the cord axis and $M_1 \sin \alpha$ about the circumferential axis. For the two plies the circumferential components cancel out. The component about the cord axis due to the two plies is

$$2M_1 \cos \alpha. \tag{8}$$

In addition to these two components, there is a third component G due to the bending of the ply axis. Since no complete theory of the problem of the bending of a circular rod subjected to *large* deformations is available, this component is calculated from the classical (small-deformation) theory by means of the formula

$$G = EI/\rho \tag{9}$$

where E is the Young's modulus, I the moment of inertia of the cross section, and ρ the radius of curvature. For E, however, the value used is not that corresponding to the unstrained state (as in the classical theory) but the effective longitudinal modulus of the rubber in the stretched state, for a small superimposed deformation. This is given by E_λ, where

$$E_\lambda = \frac{dF_1}{dl}\frac{l}{A} = \frac{dF_1}{d\lambda}\frac{\lambda}{A} \tag{10}$$

where F_1 is the tensile force on a rod of length l and area A, measured in the strained state. Putting $\tau = 0$ in Eqn (1) and differentiating, we obtain

$$\frac{dF_1}{d\lambda} = 2\pi b_0^2 \left\{ C_1 \left(1 + \frac{2}{\lambda^3}\right) + \frac{3C_2}{\lambda^4} \right\}. \tag{11}$$

This refers to a rod whose initial radius is b_0 and whose cross-sectional area A in the strained state is (from the condition of constancy of volume),

$$A = \pi b_0^2/\lambda. \tag{12}$$

Insertion of (11) and (12) in (10) gives

$$E_\lambda = 2C_1 \left(\lambda^2 + \frac{2}{\lambda}\right) + \frac{6C_2}{\lambda^2}. \tag{13}$$

This is the effective modulus in the axial direction for a rod without twist, at the extension ratio λ. For a circular rod of radius b (in the strained state) the value of I is

$$I = \pi b^4/4. \tag{14}$$

Substituting Eqns (13) and (14) in Eqn (9) we obtain for the bending moment

$$G_1 = \frac{\pi b^4}{4\rho} \left\{ 2C_1 \left(\lambda^2 + \frac{2}{\lambda}\right) + \frac{6C_2}{\lambda^2} \right\}. \tag{15}$$

This couple acts about an axis normal to the principal plane of curvature of the ply axis, i.e. about an axis inclined at an angle $\frac{1}{2}\pi - \alpha$ to the cord axis. It may therefore be resolved into a component couple about the cord axis of amount $G \sin \alpha$, and a component couple about a circumferential axis of amount $G \cos \alpha$. For the two plies the two circumferential components cancel out, and the resultant axial couple is

$$2G_1 \sin \alpha. \tag{16}$$

The total couple M_c about the cord axis due to the tensile force F_1, the torsional couple M_1 and the bending moment G_1 is therefore, from Eqns (7), (8) and (16),

$$M_c = 2M_1 \cos \alpha + 2F_1 a \sin \alpha + 2G_1 \sin \alpha. \tag{17}$$

2.4. *Corrections to above formulae.* The use of the foregoing formulae may lead to significant errors, particularly at high values of cord twist. These errors arise primarily from the curvature of the ply axis, which introduces a lateral pressure between the two plies. The effects of this lateral pressure are (*a*) to reduce the force along the ply axis required to maintain a given extension of the ply axis, and (*b*) to produce a distortion or flattening of the cross section which will affect both the torsional couple about the ply

axis and the bending moment. The distorted section may be approximated by an ellipse having major and minor semi-axes b_1 and b_2. Experimentally b_2 is taken to be equal to a (Fig. 1) while b_1 is obtained from measurements of the ply diameter in the perpendicular direction. (At the highest twists values of b_1/b_2 as high as 1·32 were obtained.)

Corrections have been worked out for (i) the effect of the lateral pressure on the tensile force F_1 along the ply axis, (ii) the effect of non-circularity of section on the torsional couple M_1, and (iii) the effect of non-circularity on the bending moment G_1. In addition an attempt has been made (iv) to correct the value of E_λ (which is involved in the calculation of G_1) for the effect of torsion. The calculations of these corrections, which are given in the Appendices, are essentially rather crude, but should be sufficient to enable an estimate of the respective orders of magnitude involved to be made. From the numerical values of the corrections, given later, it appears that the only one which is of real significance in the present experiments is (i).

3. The experimental arrangement

The experiments involved the measurement of axial tension and torsional couple in both the single rod and the two-ply 'cord'. The apparatus used for these measurements is shown diagrammatically in Fig. 2; A is a rigid mounting

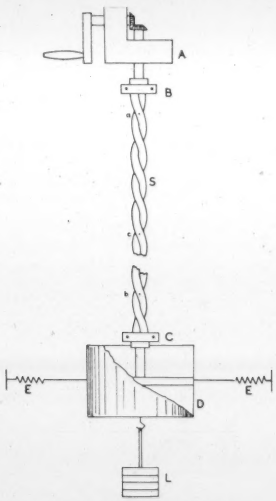


Fig. 2. Diagram of apparatus used for measurements of axial tension and torsional couple.

to which was attached the upper clamp B which could be rotated about a vertical axis. The lower clamp C was rigidly fixed to the drum D of 6 cm diameter. Rotation of this drum was limited by threads attached to the calibrated springs E. The length of the specimen S (single rod or cord) between clamps was 25 cm; measurements of strain parameters were made on a 10-cm marked length a c, by means of a low-power microscope, whilst the twist was obtained from the number of turns in the 20-cm length a b. Before mounting, the rods were stretched to twice their length and released about twenty times; measurements of the unstrained dimensions (length and diameter) were made after this preliminary treatment.

The rubber used was a commercial pure-gum vulcanized rod of diameter 5 mm and specific gravity 1·048.

4. Numerical calculations

From the measured geometrical parameters of the cord and the values of C_1 and C_2 determined experimentally, as described in the following paragraph, the components F, M and G were calculated, using the formulae given in Section 2 together with the corrections set out in the Appendices. The contributions to the total cord torque were calculated by means of Eqns (7), (8) and (16), in which F_1, M_1 and G_1 are replaced by the corrected values F, M and G. The separate components of torque thus calculated are shown in Figs 4, 5, 7 and 8, together with the total torque, obtained by summation of these components. The calculated values of axial load are given in Figs 6, 9 and 10.

The effect of the various correction terms on the components of the total torque is shown in the accompanying Table, which refers to the experiments with zero single rod torsion. It will be seen that all the correction terms increase in importance with increasing twist, and that the most serious is that due to the effect of lateral pressure on F. The corrections due to the effects of torsion in the ply and non-circularity of section respectively on the bending couple act in opposite senses and nearly cancel each other out.

5. Extension and torsion of single rod: determination of elastic constants

The experiments on the combined extension and torsion of a single rod involved the measurement of axial couple as a function of twist, at constant axial length. This necessitated adjustment of the load L for each value of twist. The relations between axial couple and torsion in the rod, for various values of the extension ratio λ, are shown in Fig. 3; these are approximately linear, in agreement with Eqn (2). The inset is included to show that this linearity is maintained up to values of torsion equal to the highest encountered in the doubled rod experiments. (This occurred in the case in which the initial single rod torsion was −7·11 rad/cm, where the ply torsion reached a value of −6·08 rad/cm with a ply extension ratio of 1·56.)

The degree of reversibility of the torque is indicated by

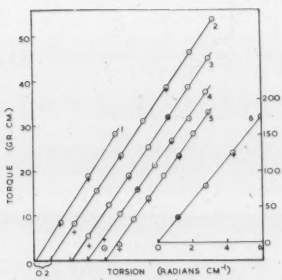


Fig. 3. Torque plotted against torsion in single rod at extension ratio (1) 1·01, (2) 1·19, (3) 1·38, (4) 1·61, (5) 1·85 and (6) inset, 1·59. The origin is displaced 0·2 rad/cm at each extension ratio. O twist increasing, + twist decreasing.

Table of corrections to the cord torque (zero single rod torsion)

Cord extension ratio, λ_c	*Cord torsion* (*rad/cm*)	*Cord torque* (*uncorrected*) (*g cm*)	*Effect of lateral pressure on* $2Fa \sin\alpha$ (*g cm*)	*Effect of non-circularity on* $2M \cos\alpha$ (*g cm*)	*Effect of non-circularity on* $2G \sin\alpha$ (*g cm*)	*Effect of torsion on* $2G \sin\alpha$ (*g cm*)	*Total correction* (*g cm*)	*Total correction* (%)
	0·62	38·31	−0·05	0·00	0·00	+0·01	− 0·04	−0·1
	1·25	85·40	−0·54	0·00	−0·21	+0·21	− 0·54	−0·6
1·01	1·87	135·40	−2·31	0·00	−1·04	+1·10	− 2·25	−1·7
	2·50	184·27	−5·83	0·00	−3·06	+2·84	− 6·05	−3·3
	3·12	219·71	−9·74	−0·49	−6·23	+5·55	−10·91	−5·0
	0·39	47·32	−0·05	0·00	0·00	0·00	− 0·05	−0·1
	0·77	90·61	−0·41	0·00	−0·03	+0·01	− 0·43	−0·5
	1·16	132·79	−1·30	0·00	−0·13	+0·05	− 1·38	−1·0
1·64	1·54	171·14	−2·75	−0·71	−0·44	+0·18	− 3·72	−2·2
	1·93	202·50	−4·52	−1·67	−0·89	+0·45	− 6·63	−3·3
	2·31	228·73	−6·54	−2·90	−1·74	+0·93	−10·25	−4·5
	2·70	255·59	−9·07	−4·24	−2·90	+1·64	−14·57	−5·7

the points obtained during untwisting of the rods, which are shown in Fig. 3.

To obtain the values of the constants C_1 and C_2 in Eqn (2) the quantities $M/\pi b_0^4\tau$, obtained from the slopes of the lines in Fig. 3, were plotted against $1/\lambda$. In accordance with Eqn (2) this should yield a straight line, whose intercept on the axis $\lambda = 1$ is equal to $C_1 + C_2$ and whose slope is C_2. Five sets of values of C_1 and C_2 were obtained. From the nature of this experiment the value of $C_1 + C_2$ can be obtained more accurately than the values of C_1 and C_2 separately. (Conversely, the individual values of C_1 and C_2 have a smaller effect on the calculated values of the stresses than does their sum $C_1 + C_2$; their lower accuracy is therefore less significant.) For $C_1 + C_2$ the values obtained were within $\pm 3\%$, while for C_1 and C_2 separately the scatter was much greater. The mean values obtained were

$$C_1 = 1{\cdot}64\,\text{kg/cm}^2,\ C_2 = 0{\cdot}62\,\text{kg/cm}^2,\ C_1 + C_2 = 2{\cdot}26\,\text{kg/cm}^2. \tag{18}$$

6. Experiments on corded rods

6.1. *Without initial twist in single rods.* The two rods were mounted parallel to each other between the clamps B and C (Fig. 2) with the reference marks on the same horizontal line. For each value of cord twist the axial load was adjusted so as to keep the axial extension ratio λ_c constant. The axial length was measured on a travelling microscope, which was also fitted with an eyepiece scale for the measurement of the cord diameter $4a$ and the larger ply diameter $2b_1$ (Fig. 1). Mean values of several measurements of these quantities along the length of the cord were used. The experiment was carried out at six values of λ_c ranging from 1·01 to 1·64. The results for the two extreme values only are reproduced here; the remaining data were intermediate, and their inclusion would not reveal any features which are not brought out by the results actually reported. The maximum cord twist employed corresponded to a cord helix angle of about 33°. The relationship between cord torque and cord torsion for these two values of axial extension is given in Figs 4 and 5, and the corresponding values of axial load in Fig. 6.

6.2. *With initial twist in single rods.* The initial twist was inserted in the single rods by clamping them side by side at one end and rotating the other ends independently through the same number of turns. The top ends were then locked together and the cord twist inserted in the usual way. Two experiments were carried out; in the first the single rod torsion was −2·21 rad/cm and the axial extension ratio 1·14, while in the second the single rod torsion was −7·11 rad/cm and the axial extension ratio 1·48. The negative sign implies that the singles twist is in the opposite sense to the positive cord twist. (The second experiment included both positive and negative cord twists.) The relation between cord torque and cord torsion is given in Figs 7 and 8, and the corresponding values of axial load in Figs 9 and 10.

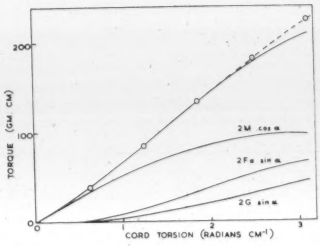


Fig. 4. Torque plotted against torsion in corded rubber rods. Rods initially without twist. —— theoretical. ○ experimental. Cord extension ratio = 1·01.

The calculated curve of Fig. 10 is carried beyond the range of the experimentally applied torsion in order to show the maximum and minimum. This calculation involved a small

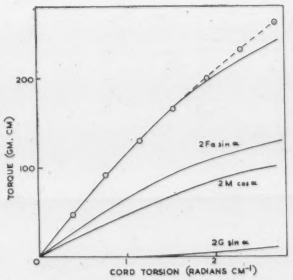


Fig. 5. Torque plotted against torsion in corded rubber rods. Rods initially without twist. —— theoretical. ○ experimental. Cord extension ratio = 1·64.

extrapolation of the measured values of a. Any possible errors in this extrapolation would not be sufficient to alter the general form of the calculated curve.

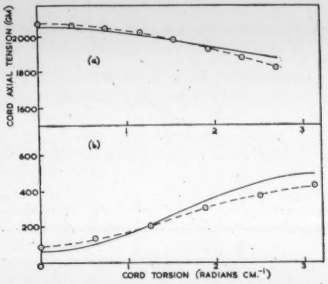


Fig. 6. Axial tension plotted against torsion in corded rubber rods. Rods initially without twist. Extension ratios, (a) 1·64 and (b) 1·01. —— theoretical. ○ experimental.

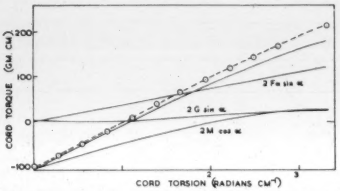


Fig. 7. Torque plotted against torsion in corded rubber rods. Rods initially twisted. —— theoretical. ○ experimental. Cord extension ratio = 1·14. Initial single rod torsion = − 2·21 rad/cm.

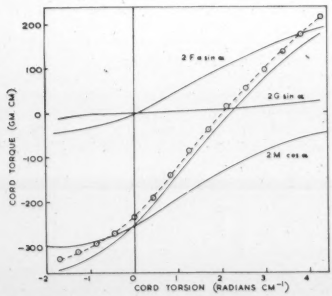


Fig. 8. Torque plotted against torsion in corded rubber rods. Rods initially twisted. —— theoretical. ○ experimental. Cord extension ratio = 1·48. Initial single rod torsion = − 7·11 rad/cm.

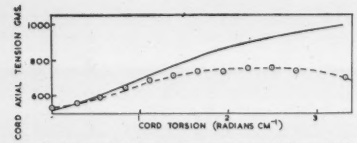


Fig. 9. Axial tension plotted against torsion in corded rubber rods. Rods initially twisted. —— theoretical. ○ experimental. Cord extension ratio = 1·14. Initial single rod torsion = − 2·21 rad/cm.

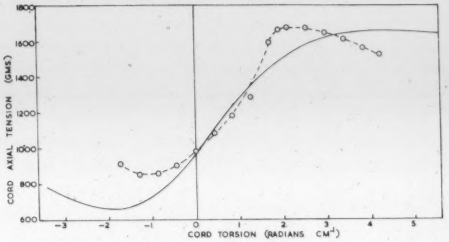


Fig. 10. Axial tension plotted against torsion in corded rubber rods. Rods initially twisted. —— theoretical. ○ experimental. Cord extension ratio = 1·48. Initial single rod torsion = − 7·11 rad/cm.

7. Comparison of calculated and observed forces

7.1. *Cord torque.* In the case of zero single-rod torsion the calculated and observed values of cord torque are in close agreement, except at the highest twist. When single rod torsion is present there is a small but consistent discrepancy over the whole range of cord torsion. This discrepancy is such that at zero cord torsion the observed couple is numerically lower than that calculated. At this point the couple is simply that due to the straight twisted rods, and the discrepancy can only be interpreted as a lack of reproducibility in the stress–strain properties of the rubber. This discrepancy, and hence also the discrepancies at other values of cord torsion, is consistent with a greater degree of stress relaxation in the experiments involving cord torsion, compared with the single-rod experiments—a not unreasonable supposition in view of the longer time taken.

7.2. *Cord axial tension.* Comparing the values of axial tension, we find fairly close agreement with theory for the case of zero single-rod torsion. For the case when single-rod torsion is present, however, the deviations are larger, particularly at high values of cord torsion. These deviations may arise either from the non-fulfilment of the assumptions upon which the theory is based, or from the admitted inadequacy of the theoretical treatment in the case when the curvature of the ply axis is large and the departure from circularity of the section considerable. It would not appear profitable at the present stage to attempt a more exhaustive analysis of the theoretical problem, but it is hoped that further work, involving multi-filament yarns, will enable a more balanced assessment of the adequacy of the present type of theory to be obtained.

Appendix 1

Correction of the tensile force in the plies due to lateral pressure

A lateral pressure exerted on the curved surface of a cylindrical rod will tend to cause an elongation in the axial direction. If the rod is subjected also to an axial tension, the effect of the lateral pressure is to reduce the axial tension required to maintain a given axial extension. The effect on the axial tension may be obtained by superimposing a negative hydrostatic pressure so as to reduce the mean lateral pressure to zero. For a material having volume incompressibility this does not affect the state of strain.

The effect will first be obtained for a curved rectangular strip; this result will then be applied to the case of a rod of elliptical section.

Rectangular strip bent into a cylinder. Consider a rectangular strip bent into the form of a cylinder having the section shown in Fig. 11. Let X be the thickness, ρ the radius of curvature of the neutral axis and 2θ the (small) angle subtended at the cylinder axis. The dimensions

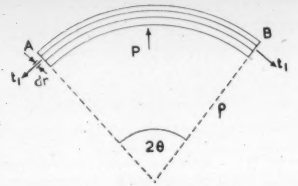


Fig. 11.

parallel to the axis will be taken as unity. The tensile stress t_1 acting on the end surface will be assumed to be constant. If p is the pressure in the radial direction, the radial force acting on a surface at the radial position r is $p(2\theta r)$. The resultant outwardly directed radial force acting on a shell of radial thickness dr is therefore

$$-2\theta\{(p+dp)(r+dr)-pr\} = -2\theta(pdr+rdp). \quad \text{(A1)}$$

For this to be in equilibrium with the sum of the radial components of the tensile forces $t_1 dr$ acting at each end, it follows that

$$2t_1\theta dr = -2\theta(pdr+rdp) \quad \text{(A2)}$$

or

$$r\frac{dp}{dr} = -t_1 - p. \quad \text{(A3)}$$

Integration subject to the boundary condition $p = 0$ when $r = \rho + \frac{1}{2}X$ gives

$$p = t_1\left(\frac{\rho+\frac{1}{2}X}{r} - 1\right). \quad \text{(A4)}$$

Eqn (A4) represents a linear rate of fall of pressure from a maximum at the inside surface to zero at the outside. The pressure P on the inside surface has the value

$$P = t_1\left(\frac{X}{\rho-\frac{1}{2}X}\right). \quad \text{(A5)}$$

The mean value of the radial pressure throughout the specimen is $\frac{1}{2}P$, which, for small curvatures ($\frac{1}{2}X \ll \rho$) may be written

$$\bar{P} = \tfrac{1}{2}P \simeq t_1 X/2\rho. \quad \text{(A6)}$$

It will now be assumed that for the purpose of calculating the effect of the lateral pressure on the tensile stress, the actual pressure distribution may be replaced by a constant pressure $\bar{P}$ equal to its mean value. (This will be true for small values of $\bar{P}$.) The resultant state then corresponds to a homogeneous stress, with principal (tensile) components t_1, t_2 and t_3 (Fig. 12) having the values

$$\begin{aligned} t_1 &= t_1 \\ t_2 &= 0 \\ t_3 &= -\bar{P}. \end{aligned} \quad \text{(A7)}$$

Superimposing a hydrostatic pressure $-\frac{1}{2}\bar{P}$, the corresponding principal stresses become

$$\begin{aligned} t_{10} &= t_1 + \tfrac{1}{2}\bar{P} \\ t_{20} &= +\tfrac{1}{2}\bar{P} \\ t_{30} &= -\tfrac{1}{2}\bar{P} \end{aligned} \quad \text{(A8)}$$

the mean lateral stress $\frac{1}{2}(t_{20}+t_{30})$ is now zero, but the state of strain is unchanged. Thus t_{10} is the tensile stress that would have to be applied, with zero mean lateral stresses, to maintain the same strained length of the specimen.

It does not follow immediately that the tensile stress t_{10} corresponds to that for a rod in simple extension. For this to be so it is necessary that the presence of the lateral stresses t_{20} and t_{30} in (A8) shall not affect the tensile stress t_{10}, for a constant extension ratio λ_1 in the direction of the rod axis. This is certainly true for small deformations, since the two lateral stresses are equal and opposite (corresponding to a shear stress). For large deformations it is no longer exactly

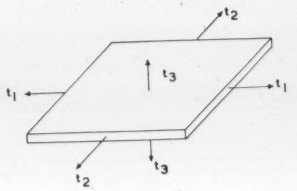


Fig. 12.

true, but it is easy to show that the effect of the transverse stresses on the tensile stress required to keep λ_1 constant is proportional to the square of the lateral strains, and hence is not important unless these are comparatively large. In the present experiments the highest ratio of major to minor axes was 1·32. With the experimental values of C_1 and C_2 this would give a maximum change in t_{10} of only 0·25%, which is negligible.

We may therefore assume that the relation between the tensile stress t_{10} as calculated for a straight rod and the actual tensile stress t_1 in the curved rod is represented by the first of Eqns (A8). Inserting the value of $\bar{P}$ from (A6), this yields

$$t_{10} = t_1\left(1+\frac{X}{4\rho}\right). \quad \text{(A9)}$$

Rod of elliptical section. Consider a rod of circular section subjected to axial extension and bent into a circular arc of radius ρ. We will assume that its section in the deformed state is an ellipse (Fig. 13) having principal semi-axes b_1 and b_2, the axis b_2 lying in the plane of curvature. As

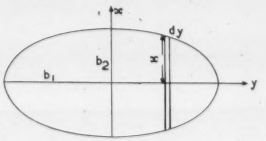


Fig. 13.

in the case of the rectangular strip, let t_{10} be the stress, and f_0 the corresponding axial force required to maintain the axial extension in the case when the mean lateral stress is zero,

and t_1 the mean stress on a strip of thickness $2x$ and width dy referred to axes coinciding with the principal axes of the ellipse. Applying the above result for the rectangular strip the tensile force df on this strip is

$$df = t_1 2x dy = t_{10} \frac{2x dy}{1 + 2x/4\rho}. \quad \text{(A10)}$$

The total force over the section is therefore

$$f = 2t_{10} \int_{-b_1}^{b_1} \frac{x dy}{1 + x/2\rho}. \quad \text{(A11)}$$

If ρ is large this is approximately

$$f = 4t_{10} \int_0^{b_1} \left(x - \frac{x^2}{2\rho}\right) dy \quad \text{(A12)}$$

$$= t_{10}\left(\pi b_1 b_2 - \frac{4}{3}\frac{b_1 b_2^2}{\rho}\right). \quad \text{(A13)}$$

Putting $f_0 = \pi b_1 b_2 t_{10}$,

$$f = f_0\left(1 - \frac{4}{3\pi}\frac{b_2}{\rho}\right). \quad \text{(A14)}$$

In this expression f_0 is the axial force when the mean lateral stress is zero, i.e. when t_{20} and t_{30} have equal and opposite values, as in (A8). It is these residual stresses, which together constitute a shear stress, that maintain the lateral strain, i.e. the ellipticity. If these lateral stresses are removed the rod returns to its original circular shape, but the tensile force (at constant axial extension) is unaffected. Hence f_0 is the force required to produce the same axial extension in a straight circular rod.

The above argument assumes that the strains associated with the lateral curvature may be dealt with on small-deformation theory. This implies that the radius of curvature is large compared with the radius of the rod, and that the ellipticity is small. In our actual experiments this condition was not always satisfied, but since the effects here considered are only brought in as a *correction* to the large-deformation theory, this imperfection should not be serious. The reason for including the elliptical, rather than the circular section in the analysis is that it shows that the significant dimension for the calculation of this correction is the minor axis b_2 rather than the mean radius $b_0/\sqrt{\lambda}$.

Appendix 2

Corrections due to non-circularity of section of the rods

(*a*) *Torsional couple M.* For a homogenous, isotropic, elastic material of elliptical cross section, the twisting couple, on the basis of the classical (small-strain theory) is (Love 1934),

$$\frac{\mu\tau\pi b_1^3 b_2^3}{b_1^2 + b_2^2} \quad \text{(A15)}$$

where b_1 and b_2 are the major and minor semi-axes of the ellipse, τ is the torsion and μ is the modulus of rigidity. For $b_1 = b_2 = r$ this reduces to the usual form

$$\tfrac{1}{2}\mu\tau\pi r^4. \quad \text{(A16)}$$

The large-strain equivalent of (A16) is

$$M_1 = \pi\tau C_1 + \frac{C_2}{\lambda} b_0^4 = \pi\tau C_1 + \frac{C_2}{\lambda}\lambda^2 r^4 \quad \text{(A17)}$$

where b_0 and r are the unstrained and strained radii, respectively. It will be assumed that the correction for non-circularity in the case of large strains involves the replacement of r^4 by the same function of b_1 and b_2 as in the classical theory. The result is then

$$M = \pi\tau C_1 + \frac{C_2}{\lambda}\lambda^2 \frac{2b_1^3 b_2^3}{b_1^2 + b_2^2} \quad \text{(A18)}$$

where b_1 and b_2 are the radii measured in the strained state. We may write

$$\frac{2\lambda^2 b_1^3 b_2^3}{b_1^2 + b_2^2} = \frac{2b_1/b_2}{1 + b_1^2/b_2^2}\lambda^2 b_1^2 b_2^2 = \frac{2b_1/b_2}{1 + b_1^2/b_2^2} b_0^4 \quad \text{(A19)}$$

since for an incompressible material $\lambda b_1 b_2 = b_0^2$. It follows from Eqns (A17) and (A18) with (A19) that

$$\frac{M}{M_1} = \frac{2b_1/b_2}{1 + b_1^2/b_2^2} = \frac{2q}{1 + q^2}. \quad \left(q = \frac{b_1}{b_2}\right) \quad \text{(A20)}$$

where M_1 is calculated from Eqn (2).

(It is assumed in the above treatment and also in paragraph (*b*) below, that the ellipticity of section is present in the unstrained state. Actually, the unstrained section is circular, and the ellipticity is produced by the forces acting on the system. In small-deformation theory this would not affect the result, but this cannot strictly be assumed in the case of large deformations.)

(*b*) *Bending couple G.* Assuming that the longitudinal modulus remains unchanged, we are concerned only with the change in shape of the cross section.

The derivation of the bending couple G_1 (Eqn (15)) was carried out on the basis of a circular cross section of radius b, calculated from the unstretched radius b_0 and the extension ratio λ. From the classical equation for small strains,

$$G = EI/\rho. \quad \text{(A21)}$$

For a circular section $I = \pi b^4/4$; the uncorrected couple G_1 is therefore

$$G_1 = E\pi b^4/4\rho. \quad \text{(A22)}$$

For an elliptical section of semi-axes b_1 and b_2, with b_2 in the plane of curvature

$$I = \tfrac{1}{4}\pi b_1 b_2^3. \quad \text{(A23)}$$

From Eqns (A21) and (A23) the corrected bending moment G is

$$G = E\pi b_1 b_2^3/4\rho. \quad \text{(A24)}$$

Also

$$\pi b^2 = \pi b_1 b_2. \quad \text{(A25)}$$

Comparing (A24) with (A22) we thus obtain

$$\frac{G}{G_1} = \frac{b_1 b_2^3}{b_1^2 b_2^2} = \frac{b_2}{b_1} = \frac{1}{q}. \quad \text{(A26)}$$

Appendix 3

Correction due to the effect of torsion on bending couple

To investigate the effect of torsion on the bending of a stretched rod, the rod may be considered to be composed of a number of cylindrical shells, whose moduli (parallel to the axis) are a function of their radii. The contribution of a given shell to the bending couple will be taken to be given by the usual small-strain formula; the total couple is then obtained by integrating with respect to the radius. The resultant mean effective modulus for bending is, of course, different from the mean effective modulus for simple extension of the twisted cylinder, since, in the case of bending, the

contribution of a shell depends not only on its cross-sectional area but also on its radius.

Modulus of shell. The modulus of a shell for a small superimposed extension may be obtained from the stored-energy function. For a shell of radius r_0 in the undeformed state subjected to an axial extension in the ratio λ, so that its radius in the strained state is $r = r_0/\lambda^{\frac{1}{2}}$, together with a torsion τ per unit length, measured in the strained state, the stored energy per unit volume is (Rivlin 1949)

$$W = C_1\left(\lambda^2 + \frac{2}{\lambda} + \lambda\tau^2 r_0^2 - 3\right) + C_2\left(2\lambda + \frac{1}{\lambda^2} + \tau^2 r_0^2 - 3\right). \tag{A27}$$

If θ is the angular rotation per unit length, referred to the unstrained length, then $\tau = \theta/\lambda$. In terms of θ the above equation becomes

$$W = C_1\left(\lambda^2 + \frac{2 + \theta^2 r_0^2}{\lambda} - 3\right) + C_2\left(2\lambda + \frac{1 + \theta^2 r_0^2}{\lambda^2} - 3\right). \tag{A28}$$

For a further extension $l_0 d\lambda$ at constant θ the change in W is equal to the work done by the applied force f, namely $f l_0 d\lambda$, where l_0 is the initial unstretched length.* Hence, if A_0 is the initial cross-sectional area,

$$f l_0 d\lambda = A_0 l_0 dW$$

or

$$f = A_0\left(\frac{dW}{d\lambda}\right)_\theta. \tag{A29}$$

The differential Young's modulus for the extension $l_0 d\lambda$ is the ratio of the increment of force per unit *strained* cross-sectional area to the increment of length per unit *strained* length, i.e.

$$E_\lambda = \frac{df}{A_0/\lambda}\Big/\frac{d\lambda}{\lambda} = \frac{1}{A_0}\left(\lambda^2 \frac{df}{d\lambda}\right)_\theta \tag{A30}$$

where A_0 ($= 2\pi r_0 dr_0$) is the unstrained area of cross section.

Combining this result with Eqns (A29)

$$E_\lambda = \lambda^2 (d^2W/d\lambda^2)_\theta. \tag{A31}$$

From (A28) we obtain

$$E_\lambda = \lambda^2\left(\frac{d^2W}{d\lambda^2}\right)_\theta = 2C_1\left(\lambda^2 + \frac{2 + \theta^2 r_0^2}{\lambda}\right) + 6C_2\frac{1 + \theta^2 r_0^2}{\lambda^2}. \tag{A32}$$

* In the case of a twisted rod, the quantity of f is not the force actually exerted on the annulus as it exists in the rod, but the force which would be exerted on this annulus if the shell were removed from its surroundings. The actual force is not related to the change in W by Eqns (A29), since work is also done by the other components of force (radial and azimuthal) present in the rod. A more fundamental analysis (in which the work done by the bending couple is related directly to the stored energy) is required fully to justify the simplified treatment given above.

Expressed in terms of τ and r this becomes

$$E_\lambda = 2C_1\left(\lambda^2 + \frac{2}{\lambda} + \lambda^2\tau^2 r^2\right) + 6C_2\left(\frac{1}{\lambda^2} + \lambda\tau^2 r^2\right). \tag{A33}$$

Integration with respect to radius. If dG is the contribution to the bending couple when the elementary shell is bent in such a way that the axis has a radius of curvature ρ, then

$$dG = E_\lambda I/\rho \tag{A34}$$

where $I = \pi r^3 dr$. The total couple for the whole rod is therefore

$$G = (\pi/\rho)\int_0^b E_\lambda r^3 dr \tag{A35}$$

where b is the radius of the rod in the deformed state.

Substituting Eqn (A33) for E_λ and integrating, we obtain

$$G = \frac{\pi b^4}{4\rho}\left\{2C_1\left(\lambda_2 + \frac{2}{\lambda}\right) + \frac{6C_2}{\lambda^2}\right\} + \frac{\pi b^6}{6\rho}\{2\lambda\tau^2(\lambda C_1 + 3C_2)\}. \tag{A36}$$

It is convenient to express this result as the sum of two terms, i.e.

$$G = G_1 + G_2 \tag{A37}$$

where

$$G_1 = \frac{\pi b^4}{4\rho}\left\{2C_1\left(\lambda^2 + \frac{2}{\lambda}\right) + \frac{6}{\lambda^2}C_2\right\} \tag{A38}$$

is the bending couple for the untwisted rod derived in the text (Eqn (15)), while

$$G_2 = (\pi b^6/6\rho)\{2\lambda\tau^2(\lambda C_1 + 3C_2)\} \tag{A39}$$

is the additional couple due to the torsion. The correction due to torsion is thus

$$\frac{G_2}{G_1} = \frac{2\lambda(\lambda C_1 + 3C_2)\tau^2 b^2}{3C_1(\lambda^2 + 2/\lambda) + 9C_2/\lambda^2}. \tag{A40}$$

It is seen that this correction increases as the *square* of the torsion.

Acknowledgment

This work forms part of a programme of research undertaken by the British Rayon Research Association.

References

LOVE, A. E. H., 1934, *The Mathematical Theory of Elasticity* (Cambridge: University Press).

RIDING, G., 1959, *J. Text. Inst.*, **50**, T425.

RIVLIN, R. S., 1949, *Phil. Trans A.*, **242**, 173.

RIVLIN, R. S., and SAUNDERS, D. W., 1951, *Phil. Trans A*, **243**, 251.

TRELOAR, L. R. G., 1956, *J. Text. Inst.*, **47**, T348.

—— 1958, *The Physics of Rubber Elasticity* (Oxford: University Press).

Demagnetization of ferromagnetic particles

by F. SMITH, M.Sc., A.Inst.P., Research Department, Simon-Carves Limited, Stockport

MS. received 29th September 1960

Abstract

The finely divided magnetic solids, encountered in dense medium coal preparation and iron-ore beneficiation plants, show a wide variation in their residual magnetism after passage through an alternating current demagnetizer. High residual magnetism is associated with considerable oscillation of the suspended particles in the alternating magnetic field. It is shown that this behaviour cannot be unambiguously correlated with the magnetic properties of the bulked material, and a test based on settling characteristics is preferred as a means of differentiating the materials.

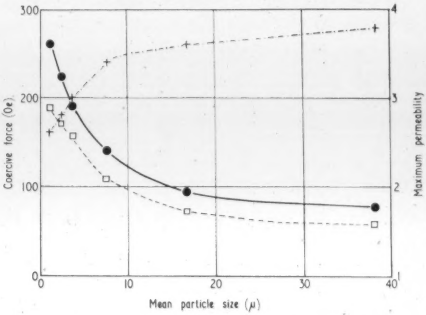
1. Introduction

An important problem in the mineral dressing industry, e.g. in coal preparation by separation in dense medium baths and in iron ore beneficiation plants employing grinding and size separation, is the demagnetization of small ferromagnetic particles in aqueous suspension to reduce their flocculation to a minimum. The normal method of demagnetization, by the application of an alternating magnetic field of sufficient strength to remove previous magnetization and the subsequent gradual reduction of this field, does not always give good results. In the sub-sieve size range there is a close balance between the magnetic couple tending to rotate the magnetized particle and the forces tending to rearrange the domain structure. Whether or not the particle rotates is dependent upon many factors, its size, shape, inertia, proximity to other particles, the medium viscosity, the magnetic characteristics and the alternating current frequency.

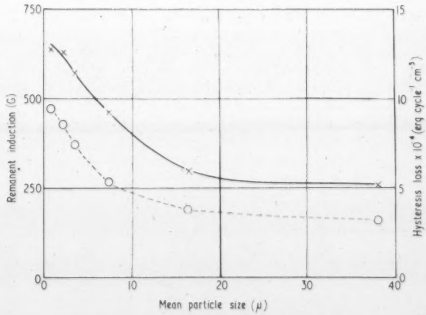
Hartig *et al.* (1951) considered the coercive force ${}_{B}H_{C}$, measured on the bulked powder, to be the main criterion for such rotation and gave a limiting value of 100 Oe, and other workers (Onstad and Foot 1954, van der Walt 1957, Williams and Hendrickson 1956) have since accepted this correlation. However, a preliminary investigation showed that the magnetic properties, including the coercive force, are a function of the particle size (Fig. 1), which would indicate that Hartig's criterion was incomplete. The work described in this paper was carried out to determine how the coercive force and the other bulk magnetic characteristics of a variety of magnetic powders, of different sizes, affects their demagnetization in suspension.

2. Measurement of the residual magnetization and demagnetization efficiency

No reliable and simple method (Morrish and Yu 1956a, b) is available for the examination of the external field associated with small particles, after incomplete demagnetization and indirect methods have to be employed. Where several magnetized particles are in close proximity they form an agglomerate, the size of which governs the settling rate in suspension. An increase in the size of the agglomerates increases the settling rate, which may therefore be used as an indirect measure of the efficiency of demagnetization.



(*a*) Coercive force and maximum permeability.
——— ● ——— = intrinsic coercive force.
– – – □ – – – = technical coercive force.
·—·—+—·—·— = maximum permeability.



(*b*) Remanent induction and hysteresis loss.
——— × ——— = remanent induction.
– – – ○ – – – = hysteresis loss.

Fig. 1. Variation of magnetic properties with mean particle size.

Hartig *et al.*, working with low concentrations of solids (3 to 5% by volume), used as a criterion the time required for the formation of a visible interface after shaking. The presence of non-magnetic particles interferes with the observation of the interface, and Williams and Hendrickson

modified the method for use with mixtures of magnetic and non-magnetic materials, by weighing the solids contained above and below a fixed level in the settling tube after a fixed time. At higher volume concentrations (greater than 10%) such suspensions generally settle with a sharply defined interface, even when they contain non-magnetic particles. For any magnetic state the settling speed of this interface is reasonably constant and reproducible, and this was used as a measure of the residual magnetism throughout the present work. The absolute values of the settling speeds do, of course, depend on the particle size distribution and the density of the sample, as well as on its magnetic history. In order to eliminate these adventitious effects, two parameters have been introduced to define the demagnetization efficiency.

(i) The 'settling rate ratio', defined as the ratio of the settling rate obtained after demagnetizing the sample in suspension to that obtained after demagnetizing the sample in the damp, compacted state.

Fig. 2 shows the settling rates of a number of materials which had been demagnetized at different specific gravities, including the damp, compacted state. Any rotation or oscillation of the particles during demagnetization gives inefficient demagnetization, and hence a higher settling rate

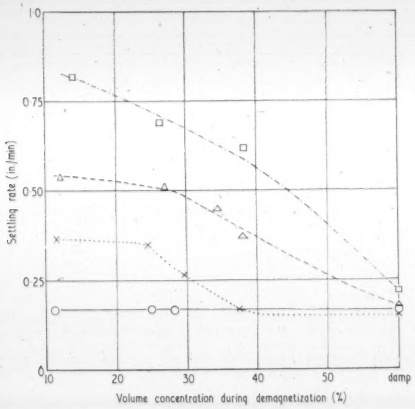


Fig. 2. Variation of settling rate of demagnetized suspension with concentration of solids during demagnetization.

- · - · - □ - · - · - = Iscor, sample 27.
- – – – △ – – – – = Shelton, sample 30.
- ····· × ····· = Consett, sample 25.
- ——— ○ ——— = P.I.C., sample 11.

than that obtained after demagnetization in the damp compacted state, where particle movement is prevented. Conversely, where no particle rotation takes place, the settling rate and the settling rate ratio are independent of the concentration of the suspension during demagnetization.

(ii) The 'equivalent magnetization field', defined as the magnetizing field required to produce, from the totally unmagnetized material, the same settling rate as that resulting from the demagnetizing treatment. This parameter eliminates the variable effect of magnetization in causing agglomeration, and hence increased settling speed, between one sample and another. This variability is shown in Fig. 3.

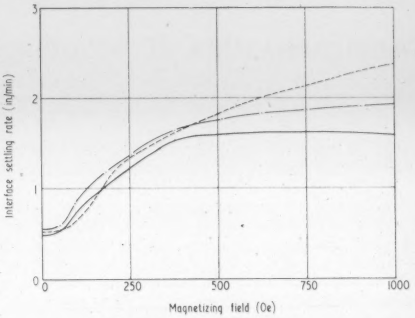


Fig. 3. Variation of interface settling rate with previously applied magnetizing field.

- – – – – – – – – = Shelton, sample 10.
- — · — · — · = Phoscor, sample 3.
- ——————— = Norwegian, sample 2.

3. Settling rate

The measurements of settling rate were made on suspensions of 12% volume concentration contained in graduated Perspex tubes 12½ in. long by 1 in. internal diameter. After each magnetization and demagnetization treatment, the suspensions were shaken vigorously to disperse the agglomerates and were allowed to settle. The time at which the interface between the suspension and supernatant water passed the different graduations was noted. The final consolidation period and a slow starting period, which occurred with some suspensions, were ignored. The rate of fall thus obtained was generally constant and reproducible. The excellent agreement between repeat measurements implies either that the vigorous shaking does not affect the residual magnetization, or that any change is completed during the first shaking. All the measurements were made with the tubes in a water bath maintained at 25 ± ½° C.

4. Magnetization and demagnetization

The suspensions were magnetized by passage through an air-cored solenoid giving a field of 50 oersteds per ampere of exciting current.

For demagnetization the suspensions, all of which had been magnetized at 500 Oe, were passed slowly through an iron-cored demagnetizer with a 2 in. gap, operating with a peak field strength of 1600 Oe at 50 c/s. This high field strength was chosen to amplify the differences arising from particle rotation.

The efficiency of the demagnetization of the damp compact, which is used as the standard throughout the present work, was assessed by measuring the settling rate after such treatment and the settling rates obtained after the following treatments:

(*a*) roasting in air at 850° C (Meerman and Oderkerken 1953), when the magnetic ferrosoferric oxide (Fe_3O_4) is converted to the weakly magnetic ferric oxide (Fe_2O_3);

(*b*) demagnetization of a frozen suspension (Hartig *et al.* 1951) of 12% volume concentration, which prevents any motion of the particles.

The settling rates given in Table 1 show that there is good agreement between the three methods.

Table 1. *Comparison of methods of demagnetization*

Medium	Interface settling rate (in/min) After demagnetization of the damp compact	After roasting	After demagnetization of the frozen suspension
Natural—Norwegian	0·36	0·37	—
Synthetic—P.I.C.	0·51	0·45	—
Natural—Norwegian	0·49	—	0·50
Blast furnace flue dust	0·38	—	0·37

5. Magnetic properties

The magnetic characteristics were measured by a modification of the Ewing Isthmus method, similar to that described by Gottschalk and Davis (1935). Briefly, the sample was placed in one of a pair of similar capsules in the gap of an electromagnet. The secondary coils wound round each capsule were balanced so that, in the absence of a sample, no galvanometer deflection was observed on changing the magnetic field. With a sample in one capsule the deflection is then proportional to the induction in the sample. Hysteresis loops were measured for maximum field strengths of 540, 1170 and 3000 Oe and the normal induction loops from 30 up to 3000 Oe.

With the exception of the intrinsic coercive force, ${}_{I}H_{C}$, the magnetic characteristics are very dependent upon the packing density (Gottschalk and Davis 1935). To compare the different materials, a standard packing density of 2·00 g/cm^3 was adopted, magnesium oxide being added as a diluent when necessary. Being non-magnetic the addition of the oxide does not add to the magnetic induction of the compact. To allow for the small unavoidable experimental variations in the actual packing density ($\pm 0{\cdot}05$ g/cm^3), a series of auxiliary measurements was made in which the packing density was varied over wide limits. The results enabled the appropriate corrections to the measured magnetic parameters to be made.

Microscopic observations, similar to those of Hartig *et al.* (1951), were also made on the behaviour of the particles under the influence of a 50 c/s alternating field.

6. Materials

The materials examined consisted of natural and synthetic magnetites, pyrites cinder extracts and blast furnace flue dust extracts (B.F.F.D.E.). They were prepared by ball milling and subsequent removal of the fraction coarser than 200 mesh B.S.S. (76 μ) and the non-magnetic material. Where possible several samples were prepared, differing in ball milling times.

To investigate the effect of chemical composition and structure, several samples of a synthetic magnetite (P.I.C.) were reduced (by hydrogen) or oxidized (by air). For the reduction the samples were placed in a thin layer in Inconel boats in a silica tube and were brought up to the required temperature of 560 $\pm$ 10° c in an inert stream of nitrogen. Hydrogen at atmospheric pressure was then passed through for the required time and the samples allowed to cool slowly in an atmosphere of nitrogen. On contact with air the reduced samples re-oxidized, sufficiently rapidly to raise the temperature to between 300 and 400° c. After re-oxidation the individual particles are probably made up of a complex matrix of the individual iron oxides and free iron, the outer layer probably being the ferroso-ferric oxide and the inner layer free iron, with ferrous oxide as the intermediate layer. Evidence for the existence of this layered structure is the further rapid re-oxidation that occurs when the individual particles are fractured by crushing or grinding.

7. Results

The variation of bulk magnetic characteristics with the particle size of the sample is shown in Fig. 1. The measurements were made on six size fractions of Norwegian magnetite separated on a Bahco air classifier. The sieved, $-53\ \mu$, sample was cut at 1·7, 2·7, 4·4, 10·5 and 22·5 μ, which gave the mean sizes of the fractions as 0·8, 2·2, 3·6, 7·5, 16·5 and 38·0 μ. It can be seen that, with the exception of the maximum permeability which shows just the reverse behaviour, the magnetic characteristic decreases with increasing particle size up to about 30 μ and then steadies off.

The principal experimental results are tabulated in Table 2, which shows the settling behaviour, the demagnetization behaviour, the bulk magnetic properties and the iron/iron oxide analysis of thirty-one materials.

In order to assess whether or not there was any correlation between the bulk magnetic properties and the demagnetization behaviour, correlation coefficients and corresponding probability levels were calculated and are given in Table 3; the samples were grouped for convenience.

8. Discussion

The qualitative microscope observations of particle chaining and rotation agree very well with the equivalent magnetization fields given in Table 2. Those media having equivalent magnetization fields of zero show, in the magnetized state, only slight rotation which ceases at relatively low alternating field strength. After demagnetization in dilute suspension no rotation is observed. For those media with high equivalent magnetization fields all the particles show rotation, both after magnetization and demagnetization in a dilute suspension.

Over the size range considered, corresponding to interface settling rates of 0·1 to 0·6 in./min at 12% volume concentration, the effect of the particle size on the equivalent magnetizing field is small.

The only significant statistical correlations, between the equivalent magnetization field or settling-rate ratio and the magnetic characteristics, that holds for all groups of media, are those for the coercive force and the maximum permeability. The correlation with the 'technical' coercive force (${}_{B}H_{C}$) suggested by Hartig *et al.* (1951) is not as good as with the 'intrinsic' (${}_{I}H_{C}$) coercive force. The fact that Hartig was working with coarser materials in dilute suspension may have some bearing upon this. Reference to Fig. 4, however, shows that even with this good correlation the coercive force cannot be used to predict with any certainty the behaviour on demagnetization. The 'natural' magnetites, i.e. excluding the blast furnace flue dust extracts and the hydrogen reduced samples, show the effect of the particle size; the solid circles, representing coarser material samples 1–7, have a lower coercive force for the same equivalent magnetizing field. That the bulk magnetic characteristics can give misleading information is readily seen by comparing samples Nos 5 and 10, Normetal pyrites and Shelton, which have similar size distributions. They have almost identical bulk magnetic characteristics (Table 2) and yet the equivalent magnetizing fields are 80 and 490 Oe respectively. After demagnetization in suspension the interface settling rates were 0·7 and

1·8 in/min, yet after efficient demagnetization in the damp compacted state, the settling rates are only 0·6 and 0·5 in/min respectively. Nevertheless, the results show that a high maximum permeability and low coercive force are, in general, associated with easy demagnetization of these small particles when in suspension. A similar correlation is known to exist between these two parameters and the ease of domain rearrangement during the magnetization of large bodies.

From their behaviour in dilute suspension the media seem to fall into two categories: (*a*) the blast furnace group, typified by extreme rotation and inefficient demagnetization in suspension; and (*b*) a natural group, including synthetic magnetite and pyrites cinder, with very much better demagnetization characteristics. The reason for the pronounced rotation of the former group is probably associated with their mode of formation, and this seems to be borne out by the experiments with reduced and oxidized synthetic magnetite. The oxidation and reduction reactions in the blast furnace flue gases will tend to produce a layered structure of oxides and free iron. The inhomogeneities will restrict the free movement of the domain walls required for demagnetization and the particles will therefore tend to take the easier path of bodily rotation with only a limited amount of domain rearrangement. The reduced P.I.C. samples, prepared so as to form such a layer structure, exhibit very high equivalent magnetization fields (Fig. 4). Where the outer layer is essentially non-magnetic, as in the oxidized P.I.C. sample, there is no apparent change in the extent of

Table 2. *Demagnetization, bulk magnetic properties and iron content of magnetic materials*

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
1	P.I.C.	0·63	1·00	0	57	46	23·5	2·62	248	2·01	4·20	130	0	42·4	21·3
2	Norwegian	0·49	1·00	0	96	76	38·7	7·00	356	2·09	3·72	190			
3	Phoscor	0·55	1·02	40	105	78	36·5	7·26	342	1·96	3·32	230	0	45·5	20·9
4	Rossington	0·61	1·13	60	115	93	41·5	10·60	428	2·08	3·58	250			
5	Normetal pyrites	0·63	1·14	80	120	75	23·0	3·50	175	1·52	2·05	280			
6	Ermelo	0·47	1·10	120	115	82	41·0	5·60	270	1·91	2·86	340			
7	Indian	0·58	1·30	145	160	94	33·3	5·17	215	1·54	1·98	350			
8	Oxidized P.I.C.	0·64	1·00	0	68	37	11·7	0·78	86	1·49	2·07	160	0·1	53·6	6·7
9	Iscor (B.F.F.D.E.)	0·58	3·72	630	190	116	51·3	8·56	295	1·54	2·07	380	3·5	52·7	9·2
10	Shelton (B.F.F.D.E.)	0·51	3·56	490	127	73	25·0	3·12	163	1·54	2·02	250	11·6	37·4	9·0
11	P.I.C.	0·17	1·00	0	77	64	31·2	4·92	292	2·06	3·70	200	0	43·7	20·9
12	Norwegian	0·40	1·00	0	127	103	55·0	14·10	510	2·12	3·60	260	0	48·5	23·6
13	Ermelo	0·33	1·00	0	115	84	36·6	6·48	309	1·94	3·06	340	0	55·3	14·6
14	Rossington	0·31	1·02	20	145	110	58·4	15·40	511	2·08	3·50	330	0	45·0	25·1
15	Phoscor	0·18	1·17	120	157	117	56·0	14·10	443	1·91	2·89	330			
16	Phoscor	0·27	1·07	55	140	108	49·3	12·90	435	1·93	3·12	300			
17	Phoscor	0·34	1·05	60	117	94	42·0	10·80	425	1·96	3·31	260			
18	Normetal pyrites	0·17	1·20	140	150	98	35·9	6·66	258	1·65	2·22	340	0	53·5	15·1
19	Tata	0·37	1·46	160	148	94	37·2	6·30	274	1·61	2·20	300	0	50·7	19·5
20	Indian	0·27	1·36	170	160	100	34·8	6·38	240	1·58	2·02	380	0	56·0	14·8
21	P.I.C. reduced in H_2 for 10 min	0·21	4·55	495	122	89	63·8	8·68	390	2·09	3·28	160	11·2	33·5	15·1
22	P.I.C. reduced in H_2 for 20 min	0·26	5·90	555	153	114	80·9	12·80	442	2·08	3·01	180	15·4	28·8	23·7
23	P.I.C. reduced in H_2 for 30 min	0·26	9·85	1000	237	171	111·1	22·70	536	2·04	2·55	500	19·0	19·7	28·5
24	P.I.C. reduced in H_2 for 60 min	0·22	14·30	1100	250	176	112·0	24·30	525	1·99	2·43	550	22·5	12·3	36·2
25	Consett (B.F.F.D.E.)	0·15	2·48	375	135	94	43·0	8·15	327	1·77	2·78	270	0·2	45·2	21·6
26	Iscor (B.F.F.D.E.)	0·10	3·63	460	202	132	59·0	11·90	352	1·68	2·20	360	4·0	50·3	13·4
27	Iscor (B.F.F.D.E.)	0·22	3·74	560	203	126	54·5	10·50	327	1·68	2·11	415			
28	Iscor (B.F.F.D.E.)	0·44	3·60	560	185	118	47·4	9·30	300	1·55	2·11	400	3·8	55·1	8·0
29	Shelton (B.F.F.D.E.)	0·12	2·53	515	120	73	27·2	3·35	184	1·53	2·20	200			
30	Shelton (B.F.F.D.E.)	0·17	3·03	425	120	76	33·8	3·88	204	1·70	2·30	260	13·0	37·0	13·8
31	Shelton (B.F.F.D.E.)	0·23	3·17	660	132	83	29·4	3·92	191	1·60	2·07	260	14·0	38·6	8·1

(1) Sample number; (2) material; (3) demagnetized damp settling rate (in/min); (4) settling rate ratio*; (5) equivalent magnetizing field (Oe); (6)–(10) 3000 Oe: (6) coercive force, ${}_IH_C$, (7) coercive force ${}_BH_C$, (8) hysteresis loss $\times 10^{-3}$ (erg cycle^{-1} cm^{-3}), (9) maximum energy product $\times 10^{-3}$(GOe), (10) remanent induction, 4πI (G); (11) permeability at 1600 Oe; (12)–(13) maximum permeability: (12) value, (13) field (Oe); (14)–(16) iron content (%): (14) free, (15) ferric, (16) ferrous.

Note.—The samples having the same name but different sample number, differ in the length of time of grinding, and hence in particle size. This is reflected in the demagnetized 'damp' settling rate; the larger this value the larger the particle size.

* The ratio of the settling rate after demagnetization at 12·5% volume concentration to that after demagnetization in the 'damp' state.

Table 3. *Correlation coefficients (r) and probability levels (P) of the magnetic characteristics with the equivalent magnetizing field and settling-rate ratio*

Sample numbers	Correlation with	Coercive force: Intrinsic $_{I}H_{C}$ r	P	Coercive force: Technical $_{B}H_{C}$ r	P	Permeability: Maximum r	P	Permeability: At 1600 Oe r	P	Remanent induction r	P	Hysteresis loss r	P	Maximum energy product r	P
1 to 7	Equiv. mag. field	+0·86	0·02	+0·67	0·10	−0·84	0·02	−0·69	0·09	−0·42	—	+0·14	—	−0·03	—
	Settling-rate ratio	+0·89	0·008	+0·67	0·10	−0·81	0·03	−0·75	0·05	−0·39	—	+0·00	—	+0·01	—
1 to 10	Equiv. mag. field	+0·75	0·02	+0·59	0·08	−0·54	0·10	−0·52	—	−0·10	—	+0·40	—	+0·02	—
	Settling-rate ratio	+0·64	0·05	+0·51	—	−0·50	—	−0·49	—	−0·14	—	+0·29	—	+0·10	—
11 to 20	Equiv. mag. field	+0·73	0·02	+0·33	—	−0·92	0·001	−0·92	0·001	−0·52	—	−0·26	—	−0·29	—
	Settling-rate ratio	+0·62	0·06	+0·18	—	−0·90	0·001	−0·53	—	−0·58	0·08	−0·34	—	−0·40	—
11 to 24	Equiv. mag. field	+0·85	0·001	+0·84	0·001	−0·38	—	+0·21	—	+0·47	0·09	+0·91	0·001	+0·73	0·003
	Settling-rate ratio	+0·84	0·001	+0·85	0·001	−0·30	—	+0·27	—	+0·52	0·06	+0·92	0·001	+0·77	0·002
11 to 31	Equiv. mag. field	+0·72	0·001	+0·63	0·002	−0·50	0·02	−0·12	—	+0·12	—	+0·65	0·002	+0·43	0·05
	Settling-rate ratio	+0·68	0·001	+0·77	0·001	−0·24	—	+0·20	—	+0·41	0·07	+0·85	0·001	+0·68	0·001

The dashes indicate a probability level of greater than 0·10.

particle rotation and no apparent effect on the domain rearrangement of the central magnetic portion.

Comparing the oxidized sample with the original P.I.C. sample, it is obvious that there have been marked changes in the bulk magnetic characteristics which have not affected the demagnetization behaviour. In this case the changes have acted to reduce the percentage of magnetic material present in the particle. Similar effects can be seen within the natural group; the high values of permeability for Norwegian magnetite compared with the much lower values for the Indian magnetite can be associated with the decreased magnetite content in the Indian sample. The coercive force, which is found by experiment to be unaffected by dilution, is therefore the only available criterion of the magnetic behaviour, and it is not surprising that the correlation with demagnetization efficiency is better for this property than the others.

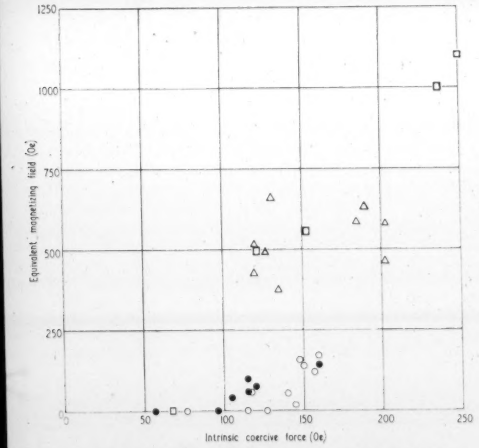


Fig. 4. Variation of equivalent magnetizing field and intrinsic coercive force.

- ● natural and synthetic magnetites, samples 1 to 7.
- ○ natural and synthetic magnetites, samples 11 to 20.
- △ blast furnace flue dust extracts.
- □ oxidized and reduced synthetic magnetite.

9. Conclusions

Although a low coercive force and high permeability are desirable, there does not seem to be any simple correlation between the bulk magnetic characteristics of all media and their behaviour on demagnetization in suspension. The measurement of the settling rate after demagnetization in suspension and in the damp compacted state gives more information about the ease of demagnetization, and this technique is now being used in the author's laboratory for the assessment of magnetic media.

Acknowledgments

The author is indebted to the Director of Research and the Management Board of Simon-Carves Ltd. for permission to publish this paper. Grateful acknowledgment is made to Messrs. D. Merrell, D. Robinson, J. M. Ormerod and D. Schofield for assistance with the experimental work and to Mr. W. Bostock and Dr. L. Cohen for helpful discussion.

References

GOTTSCHALK, V. H., and DAVIS, C. W., 1935, *U.S. Bureau of Mines, R.I.*, 3268.

HARTIG, H. E., ONSTAD, N. I., and FOOT, N. S., 1951, *Information Circular No.* 7 (University of Minnesota: Mines Experimental Station).

MEERMAN, P. G., and ODERKERKEN, H. J., 1953, *Fuel*, **32**, 62.

MORRISH, A. H., and YU, S. P., 1956a, *Rev. Sci. Instrum.*, **27**, 9.

—— 1956b, *Phys. Rev.*, **102**, 670.

ONSTAD, N. I., and FOOT, N. S., 1954, *U.S. Patent* 2,678,130.

VAN DER WALT, P. J., 1957, *Symposium on Coal Preparation* (University of Leeds).

WILLIAMS, M. F., and HENDRICKSON, L. G., 1956, *Mining Eng.*, **8**, 201.

Fields external to open-structure magnetic devices represented by ellipsoid or spheroid

by H. CHANG, Ph.D., IBM Research Center, Yorktown Heights. New York, U.S.A.

MS. received 8th August 1960

Abstract

The mathematical expressions for the field intensities external to spheroids or an ellipsoid with uniform magnetization are given. They are useful in predicting fields external to (a) open-structure magnetic devices such as metallic or ferrite films, twistors or splinters, and (b) spheroidal and ellipsoidal samples which are often used in magnetic measurements.

Curves are plotted to indicate the spatial variation of field intensity with the dimensional ratio of a spheroid as a parameter. In the region near the spheroid, field intensity is normalized against field intensity at the boundary. In the region far from the spheroid, field intensity is normalized against that of a dipole of the same dipole strength as the spheroid.

1. Introduction

Ferrite toroids and miltipaths are magnetic structures in which the magnetic flux lines complete their paths entirely within the low-reluctance ferromagnetic materials. With the advent of open-structure devices such as thin films and twistors, in which the ferromagnetic materials constitute only part of the flux path and flux lines close their paths through the air, the field external to these structures have to be considered in the interests of design and applications.

In order to stabilize the magnetic state, both for thin films and twistors, the geometries are chosen so as to minimize the demagnetizing fields. Field distributions for these devices have been analysed by the author (Chang and Milnes 1959, 1960). Recent interest in ferrite films and splinters calls for an analysis of the fields external to such structures, which is the purpose of this paper.

Spheres, oblate or prolate spheroids, or general ellipsoids are used to approximate the various structures. Arbitrarily oriented but uniform magnetization is assumed in the magnetic bodies. Field calculations are carried out for various geometrical dimensions, and normalized curves are plotted.

Various authors (Hunt 1954, Peake and Davy 1953, Stoner 1945) have treated the fields associated with an ellipsoid with uniform magnetization. In Stoner's paper expressions for the internal fields (demagnetizing fields) for an ellipsoid with uniform magnetization are given. Peake and Davy obtained expressions for external field along axes of symmetry, and Hunt derived the flux linkage with a search coil produced by a coaxial uniformly magnetized prolate spheroid.

The present paper, in addition to presenting expressions in terms of elliptical integrals for external fields of an ellipsoid with arbitrarily oriented uniform magnetization, also gives the formulae for potentials and field intensities derived in the prolate and the oblate spheroidal coordinates, instead of from the general ellipsoidal case as done by Peake and Davy and Hunt. Simple and exact equations are obtained for the external field at any point in space. Formulae for transforming rectangular into spheroidal coordinates are also given.

2. Spheroidal or ellipsoidal representation

The detailed mathematical derivations are too lengthy for presentation here. Readers interested in these derivations may consult Chang (1959). However, a discussion of the geometrical models assumed is given below.

A prolate spheroidal surface is generated by rotating an ellipse around its major axis, while an oblate spheroidal surface is generated by rotating an ellipse around its minor axis.

A general ellipsoid is represented by

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}+\frac{Z^2}{c^2}=1 \qquad (2.1)$$

where $a \geqslant b \geqslant c$ are the three semi-principal axes. When $a = b$, the ellipsoid degenerates into an oblate spheroid. When $b = c$, the ellipsoid degenerates into a prolate spheroid. When $a = b = c$, the ellipsoid becomes a sphere. The above geometries are sketched in Fig. 1. For various values assumed

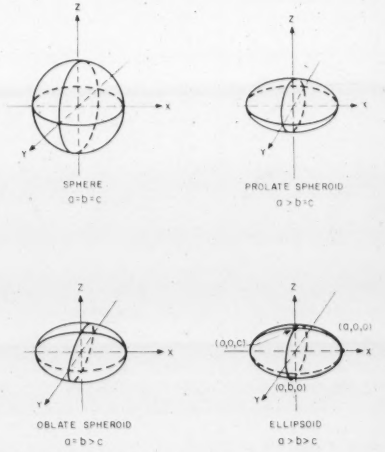


Fig. 1. Sphere, prolate and oblate spheroids and ellipsoid.

for the ratios b/a and c/a, the ellipsoid may stretch into a rod or prostrate into a disk or assume any of a range of intermediate geometries. Thus it provides an excellent model for studying the effects of geometry.

Spheroids and ellipsoids are chosen to approximate the devices, since these geometries are well suited for analysis in spheroidal or ellipsoidal coordinates. As to the justification for the above approximations, the following explanation can be offered. Assuming that the field inside a magnetic body of any geometry is known, one can easily find the external field at the boundary by applying the two boundary conditions of continuities of tangential field intensity and normal flux density. Since

$$\mathbf{B}_i = \mu_0(\mathbf{H}_i + \mathbf{M}) \tag{2.2}$$

and

$$\mathbf{B}_e = \mu_0 \mathbf{H}_e \tag{2.3}$$

where $\mathbf{B}_i$ is the flux density inside the ferromagnetic body, μ_0 the permeability of free space, $\mathbf{H}_i$ the field intensity inside the ferromagnetic body, $\mathbf{M}$ the magnetization, $\mathbf{B}_e$ the flux density external to the ferromagnetic body, and $\mathbf{H}_e$ the field intensity external to the ferromagnetic body, then the condition of continuity of tangential field intensity is expressed as

$$\mathbf{H}_{it} = \mathbf{H}_{et} \tag{2.4}$$

and the condition of continuity of normal flux density is expressed as

$$B_{in} = B_{en} \tag{2.5}$$

or

$$H_{in} + M_n = H_{en}. \tag{2.6}$$

In Eqns (2.4)–(2.6), t and n indicate tangential and normal respectively.

As the field is a continuous function in space except at the boundary, the external field distribution found from the boundary conditions depicts the field in the immediate vicinity of the ferromagnetic body as well. At a distance much larger than the dimensions of the body, the ferromagnetic body acts as a magnetic dipole in producing magnetic field there. Its geometrical shape does not affect the field there at all except in determining the magnetic dipole strength, which can be found from

$$\mathbf{m} = \int_v \mathbf{M} dV \tag{2.7}$$

where $\mathbf{m}$ is the magnetic dipole moment, V the volume of ferromagnetic body, and $\mathbf{M}$ the magnetization. The magnetostatic potential at a distance $\mathbf{r}$ from the ferromagnetic body is

$$P = (\mathbf{m} \cdot \mathbf{r})/4\pi r^3 \tag{2.8}$$

from which the field intensity is found as

$$\mathbf{H} = -\nabla P = \frac{3(\mathbf{m}\cdot\mathbf{r})}{4\pi r^5}\mathbf{r} - \frac{\mathbf{m}}{4\pi r^3}. \tag{2.9}$$

Now it remains to find the field in the intermediate region. Although the field there does depend on the geometry, the dependence is not as critical as in the immediate vicinity. Hence, a close geometrical approximation of magnetization distribution (such as an oblate spheroid for a thin circular cylindrical film) should yield results close to that of the actual distribution.

3. Fields external to spheres, spheroids or ellipsoids with uniform magnetization

The potentials due to a spheroid (oblate or prolate) or an ellipsoid of uniform magnetization are given by Chang, and that for a sphere is given by Smythe. In the following, the potential, the field intensities derived therefrom and equations for transforming rectangular coordinates into spherical, prolate or oblate spheroidal or ellipsoidal coordinates are summarized.

(i) *Sphere*

For a sphere with radius a, and uniform magnetization M_x in the X-direction, the magnetostatic potential is

$$P = \frac{M_x a^3}{3} \frac{X}{(X^2 + Y^2 + Z^2)^{3/2}}. \tag{3.1}$$

The field intensity is

$$\begin{aligned}\mathbf{H} &= H_x i_x + H_y i_y + H_z i_z \\ &= -\nabla P \\ &= -\left(\frac{\partial P}{\partial X} i_x + \frac{\partial P}{\partial Y} i_y + \frac{\partial P}{\partial Z} i_z\right).\end{aligned} \tag{3.2}$$

Hence

$$\begin{aligned} H_x &= M_x a^3 \frac{2X^2 - Y^2 - Z^2}{3(X^2 + Y^2 + Z^2)^{5/2}} \\ H_y &= M_x a^3 \frac{XY}{(X^2 + Y^2 + Z^2)^{5/2}} \\ H_z &= M_x a^3 \frac{XZ}{(X^2 + Y^2 + Z^2)^{5/2}}. \end{aligned} \tag{3.3}$$

(ii) *Prolate spheroid with uniform magnetization along the major axis*

$$P_e = M_x \frac{ab^2}{f^2} \xi \left(\frac{\eta}{2} \ln \frac{\eta + 1}{\eta - 1} - 1\right) \tag{3.4}$$

where

P_e = magnetostatic potential external to the spheroid due to magnetization,
M_x = uniform magnetization along the major axis,
a, b = major, minor axes of the prolate spheroid

$$f = (a^2 - b^2)^{\frac{1}{2}} \tag{3.5}$$

$$\eta = +\left(\tfrac{1}{2}\left[(1 + x^2 + y^2 + z^2) + \{(1 + x^2 + y^2 + z^2)^2 - 4x^2\}^{\frac{1}{2}}\right]\right)^{\frac{1}{2}} \tag{3.6}$$

$$\xi = x/\eta \tag{3.7}$$

$$\theta = \arctan z/y \tag{3.8}$$

$$x, y, z = X/f,\ Y/f,\ Z/f. \tag{3.9}$$

The rectangular components of field intensity are found as

$$\frac{H_x}{M_x ab^2/f^3} = -\left(\frac{1}{2} \ln \frac{\eta + 1}{\eta - 1} - \frac{\eta}{\eta^2 - \xi^2}\right) \tag{3.10}$$

$$\frac{H_y}{M_x ab^2/f^3} = \frac{\xi(1 - \xi^2)^{\frac{1}{2}}}{(\eta^2 - 1)^{\frac{1}{2}}(\eta^2 - \xi^2)} \cos\theta \tag{3.11}$$

$$\frac{H_z}{M_x ab^2/f^3} = \frac{\xi(1 - \xi^2)^{\frac{1}{2}}}{(\eta^2 - 1)^{\frac{1}{2}}(\eta^2 - \xi^2)} \sin\theta. \tag{3.12}$$

(iii) *Prolate spheroid with uniform magnetization along a minor axis*

$$P_e = -\frac{M_y}{2} \frac{ab^2}{f^2} (1 - \xi^2)^{\frac{1}{2}}(\eta^2 - 1)^{\frac{1}{2}} \left(\frac{1}{2} \ln \frac{\eta + 1}{\eta - 1} - \frac{\eta}{\eta^2 - 1}\right) \cos\theta \tag{3.13}$$

where M_y = uniform magnetization along a minor axis (Y-axis) and the other symbols are defined in (ii).

The rectangular components of field intensity are found as

$$\frac{H_x}{M_y ab^2/f^3} = \frac{\xi(1-\xi^2)^{\frac{1}{2}}}{(\eta^2-\xi^2)(\eta^2-1)^{\frac{1}{2}}}\cos\theta \tag{3.14}$$

$$\frac{H_y}{M_y ab^2/f^3} = \frac{1}{2}\left(\frac{1}{2}\ln\frac{\eta+1}{\eta-1} - \frac{\eta}{\eta^2-1}\right) + \frac{\eta(1-\xi^2)}{(\eta^2-\xi^2)(\eta^2-1)}\cos^2\theta \tag{3.15}$$

$$\frac{H_z}{M_y ab^2/f^3} = \frac{(1-\xi^2)\eta}{(\eta^2-\xi^2)(\eta^2-1)}\cos\theta\sin\theta. \tag{3.16}$$

(iv) *Oblate spheroid with uniform magnetization along the minor axis*

$$P_e = -M_z\frac{ca^2}{f^2}\xi(\zeta\cot^{-1}\zeta - 1) \tag{3·17}$$

where P_e = magnetostatic potential external to the spheroid due to magnetization, M_z = uniform magnetization along the minor (Z) axis, a, c = major, minor axes of the oblate spheroid

$$f = (a^2-c^2)^{\frac{1}{2}}. \tag{3.18}$$

$$\zeta = +\left(\tfrac{1}{2}\left[(x^2+y^2+z^2-1) + \{(x^2+y^2+z^2-1)^2+4z^2\}^{\frac{1}{2}}\right]\right)^{\frac{1}{2}} \tag{3.19}$$

$$\xi = z/\zeta \tag{3.20}$$

$$\theta = \arctan y/z \tag{3.21}$$

$$x, y, z = X/f, Y/f, Z/f. \tag{3.22}$$

The rectangular components of field intensity are found as

$$\frac{H_x}{M_z a^2 c/f^3} = \frac{\xi(1-\xi^2)^{\frac{1}{2}}}{(1+\zeta^2)^{\frac{1}{2}}(\zeta^2+\xi^2)}\cos\theta \tag{3.23}$$

$$\frac{H_y}{M_z a^2 c/f^3} = \frac{\xi(1-\xi^2)^{\frac{1}{2}}}{(1+\zeta^2)^{\frac{1}{2}}(\zeta^2+\xi^2)}\sin\theta \tag{3.24}$$

$$\frac{H_z}{M_z a^2 c/f^3} = \cot^{-1}\zeta - \frac{\zeta}{\zeta^2+\xi^2}. \tag{3.25}$$

(v) *Oblate spheroid with uniform magnetization along a major axis*

$$P_e = \frac{M_x}{2}\frac{ca^2}{f^2}(1-\xi^2)^{\frac{1}{2}}(1+\zeta^2)^{\frac{1}{2}}\left(\cot^{-1}\zeta - \frac{\zeta}{1+\zeta^2}\right)\cos\theta \tag{3.26}$$

where M_x = uniform magnetization along a major axis (X-axis) and the other symbols are defined in (iv).

The rectangular components of the field intensity are found as

$$\frac{H_x}{M_x a^2 c/f^3} = -\left\{\frac{1}{2}\left(\cot^{-1}\zeta - \frac{\zeta}{1+\zeta^2}\right) + \frac{\zeta(1-\xi^2)}{(1+\zeta^2)(\zeta^2+\xi^2)}\cos^2\theta\right\} \tag{3.27}$$

$$\frac{H_y}{M_x a^2 c/f^3} = \frac{\zeta(1-\xi^2)}{(1+\zeta^2)(\zeta^2+\xi^2)}\cos\theta\sin\theta \tag{3·28}$$

$$\frac{H_z}{M_x a^2 c/f^3} = \frac{\xi(1-\xi^2)^{\frac{1}{2}}}{(1+\zeta^2)^{\frac{1}{2}}(\zeta^2+\xi^2)}\cos\theta. \tag{3.29}$$

(vi) *General ellipsoid with uniform magnetization*

$$\mathbf{M} = M_x i_x + M_y i_y + M_z i_z \tag{3.30}$$

$$P = \sum_{t=x,y,z}\frac{abcM_t t}{2}\int_U^\infty \frac{ds}{(s+k^2)R_s} \tag{3.31}$$

where $(k, t) = (a, X), (b, Y)$, or (c, Z).

U defines an ellipsoidal surface:

$$\frac{X^2}{a^2+U} + \frac{Y^2}{b^2+U} + \frac{Z^2}{c^2+U} = 1, \infty > U > -c^2$$

$$R_s = \{(s-a^2)(s-b^2)(s-c^2)\}^{\frac{1}{2}}$$

$$H = -\nabla P = -\left(i_x\frac{\partial P}{\partial X} + i_y\frac{\partial P}{\partial Y} + i_z\frac{\partial P}{\partial Z}\right)$$

$$= -i_x\left\{\frac{abcM_x}{2}\int_U^\infty\frac{dU}{(U+a^2)R_u} - \frac{abcM_x}{2}\frac{X}{(U+a^2)R_u}\frac{\partial U}{\partial X} - \frac{abcM_y}{2}\frac{Y}{(U+b^2)R_u}\frac{\partial U}{\partial X} - \frac{abcM_z}{2}\frac{Z}{(U+c^2)R_u}\frac{\partial U}{\partial X}\right\}$$

$$-i_y\left\{-\frac{abcM_x}{2}\frac{X}{(U+a^2)R_u}\frac{\partial U}{\partial Y} + \frac{abcM_y}{2}\int_U^\infty\frac{dU}{(U+b^2)R_u} - \frac{abcM_y}{2}\frac{Y}{(U+b^2)R_u}\frac{\partial U}{\partial Y} - \frac{abcM_z}{2}\frac{Z}{(U+c^2)R_u}\frac{\partial U}{\partial Y}\right\}$$

$$-i_z\left\{-\frac{abcM_x}{2}\frac{X}{(U+a^2)R_u}\frac{\partial U}{\partial Z} - \frac{abcM_y}{2}\frac{Y}{(U+b^2)R_u}\frac{\partial U}{\partial Z} + \frac{abcM_z}{2}\int_U^\infty\frac{dU}{(U+c^2)R_u} - \frac{abcM_z}{2}\frac{Z}{(U+c^2)R_u}\frac{\partial U}{\partial Z}\right\}. \tag{3.32}$$

From equations defining the ellipsoidal coordinate systems, the partial derivatives in the above equation can be found as

$$\frac{\partial U}{\partial t} = \frac{2t}{K^2+U}\Big/\left\{\frac{X^2}{(a^2+U)^2} + \frac{Y^2}{(b^2+U)^2} + \frac{Z^2}{(c^2+U)^2}\right\}$$

where $(k, t) = (a, X), (b, Y)$, or (c, Z). (3.33)

The integrals in Eqn (3.32) can be reduced to incomplete elliptic integrals by the following transformation:

$$q^2 = (a^2-c^2)/(a^2+s). \tag{3.34}$$

The integral $\displaystyle\int_s^\infty \frac{ds}{(s+k^2)R_s}$

where $R_s = \{(s+a^2)(s+b^2)(s+c^2)\}^{\frac{1}{2}}$, $K = a, b$, or c (3.35)

will result in different forms for different k.

Case 1. $K = a$.

$$\int_U^\infty \frac{dU}{(U+a^2)\{(U+a^2)(U+b^2)(U+c^2)\}^{\frac{1}{2}}} = \frac{2}{(a^2-c^2)^{3/2}} \int_{q=0}^{p} \frac{q^2 dq}{\{(1-\nu^2q^2)(1-q^2)\}^{\frac{1}{2}}} = \frac{2}{(a^2-c^2)^{\frac{1}{2}}(a^2-b^2)} \{F(\nu,\psi) - E(\nu,\psi)\} \tag{3.36}$$

where

$$\nu^2 = (a^2-b^2)/(a^2-c^2)$$
$$p^2 = (a^2-c^2)/(a^2+U)$$
$$p = \sin\psi$$

$$F(\nu,\psi) = \int_0^{\psi} \frac{d\psi}{(1-\nu^2\sin^2\psi)^{\frac{1}{2}}} = \int_0^{p} \frac{dq}{(1-q^2)^{\frac{1}{2}}(1-\nu^2q^2)^{\frac{1}{2}}}$$

= incomplete elliptic integral of the first kind.

$$E(\nu,\psi) = \int_0^{\psi} (1-\nu^2\sin^2\psi)^{\frac{1}{2}} d\psi = \int_0^{p} \frac{(1-\nu^2q^2)^{\frac{1}{2}} dq}{(1-q^2)^{\frac{1}{2}}}$$

= incomplete elliptic integral of the second kind.

Case 2. $K = b$.

$$\int_U^\infty \frac{dU}{(U+b^2)\{(U+a^2)(U+b^2)(U+c^2)\}^{\frac{1}{2}}} = \frac{2}{(a^2-c^2)^{3/2}} \int_{q=0}^{p} \frac{q^2 dq}{(1-\nu^2q^2)^{3/2}(1-q^2)^{\frac{1}{2}}}$$
$$= -\frac{2(a^2-c^2)^{\frac{1}{2}}}{(b^2-c^2)(a^2-b^2)} [F(\nu,\psi) - E(\nu,\psi)] + \frac{2}{(a^2-c^2)^{\frac{1}{2}}(b^2-c^2)} F(\nu,\psi) - \frac{2}{b^2-c^2} \left\{\frac{U+c^2}{(U+a^2)(U+b^2)}\right\}^{\frac{1}{2}}. \tag{3.37}$$

Case 3. $K = c$.

$$\int_U^\infty \frac{dU}{(U+c^2)\{(U+a^2)(U+b^2)(U+c^2)\}^{\frac{1}{2}}} = \frac{2}{(a^2-c^2)^{3/2}} \int_{q=0}^{p} \frac{q^2 dq}{(1-q^2)^{3/2}(1-\nu^2q^2)^{\frac{1}{2}}}$$
$$= -\frac{2}{(a^2-c^2)^{\frac{1}{2}}(b^2-c^2)} E(\nu,\psi) + \frac{2}{(b^2-c^2)} \left\{\frac{U+b^2}{(U+a^2)(U+c^2)}\right\}^{\frac{1}{2}}. \tag{3.38}$$

4. Conclusions

(*a*) As discussed in Section 2, at a distance from the ferromagnetic body, the field distribution should be identical to that of a dipole with equal strength (Eqn (7)). For a uniformly magnetized sphere, the external field is identical to that of an equivalent dipole. For spheroids, the difference between actual field and equivalent dipole field is less than 10% beyond 3·5 times major axial (*a*) lengths, and less than 1% beyond 10*a*. Fig. 2 illustrates the above statement.

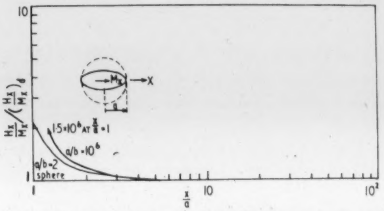


Fig. 2. H_x on X-axis for oblate spheroid, with $\mathbf{M} = M_x i_x$. (The far field is normalized against that of a dipole.)

(*b*) The larger the dimensional ratio a/b or a/c, the faster the field intensity diminishes with increasing distance from the boundary (see Fig. 3). Thus, thinner films or longer splinters tend to confine flux lines to a smaller region.

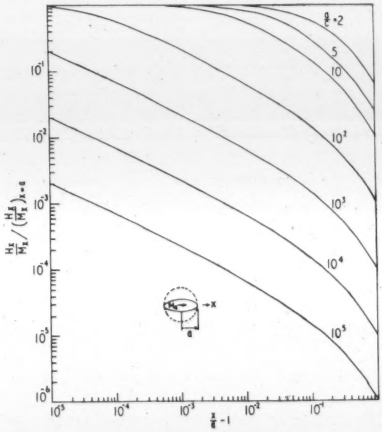


Fig. 3. H_x on X-axis for oblate spheroid, with $\mathbf{M} = M_x i_x$. (The near field is normalized against that at $x = a$.)

(*c*) To immerse a sphere, a spheroid or an ellipsoid with constant permeability μ in a uniform field results in uniform induced magnetization in the body. Thus the formulae in Section 4 can be used after induced magnetization is found:

$$M_t = H_{0t} \Big/ \left(\frac{\mu_0}{1-\mu_0} + N_t\right) \tag{4.1}$$

when one principal axis of the ellipsoid or the spheroid coincides with $t(=x, y \text{ or } z)$ axis, H_{0t} is the original field component and N_t the demagnetizing factor of the ellipsoid in the t-direction.

References

CHANG, H., 1959, *Dissertation* (Carnegie Institute of Technology, Pittsburgh).

CHANG, H., and MILNES, A. G., 1959, *Inst. Radio Engrs Trans*, **EC–8** (4), 458; 1960, **EC–9** (2), 199.

HUNT, G. H., 1954, *Brit. J. Appl. Phys.*, **5**, 260.

PEAKE, H. J., and DAVY, N., 1953, *Brit. J. Appl. Phys.*, **4**, 207.

SMYTHE, W. R., 1950, *Static and Dynamic Electricity* (New York: McGraw-Hill), p. 140.

STONER, E. C., 1945, *Phil. Mag.*, **36**, 803.

Significance of the properties of materials in the packing of real spherical particles

by J. C. MACRAE, B.Sc., F.R.I.C., M.Inst.F., and W. A. GRAY, B.Sc., Ph.D.,* The Houldsworth School of Applied Science, University of Leeds

MS. received 28th September 1960, in revised form 7th December 1960

Abstract

Measurements of the packing of spherical particles poured into a cylindrical container show that the intensity of deposition, the height of drop and the elasticity of the particle material influence the packing equally. Minor departures from perfect sphericity tend to mask the influence of elasticity.

Optimum packing of poured spherical particles occurs when the energy increment imparted to the bed by each impacting particle exceeds a critical value, the rate of renewed application of increments is above a critical value and the intensity of deposition is below a critical level. Within a critical range of deposition conditions, sufficient energy may be imparted to spherical particles to activate a process of ordered packing, initiated at a wall. These conditions are fulfilled when 0·125 in. diameter spheres fill a 3 in. diameter glass cylinder at 1400 particles per second with impact velocities greater than $120/\{\frac{1}{2}(1 + e_m)\}^{1/2}$ *in/s,* e_m *being the resilience of the particle material.*

Determinations of wall effect show that these conclusions and the derived mechanism of packing are valid for the packing of an infinitely extensive bed.

1. Introduction

THE problems associated with the behaviour of masses of solid particles are common to virtually every industry, whether the specific objective be the achievement of a dense packing or the establishing and maintenance of a freely flowing condition. Consequently, the study of the packing of granular solids has been almost continuous for many years. Despite this, there has not appeared a body of generally applicable quantitative theory helpful in the resolution of unfamiliar problems. Researches have tended to be concerned either with theoretical, unreal situations or to be entirely empirical. The former are mainly concerned with the geometry of simplified packing situations, although restricted supporting experimentation is not uncommon, and are principally the study of systems of perfect spheres. The latter appear to assume that the shapes and mechanical properties of particles exert such a powerful and unresolvable influence upon their packing behaviour that only naturally occurring, or industrially used, materials can fruitfully be used in experiments.

It is commonly appreciated that ideal spheres can be packed in layers of various types, definable by the angle of intersection of the sets of rows in a layer, which may have any value between 60° and 90°. Each kind of packing of spheres inevitably is associated with voids of characteristic shape and has a characteristic porosity and co-ordination number (or number of spheres touching every sphere). These two features together, but not separately, define a systematic packing of spheres.

The mode of packing of relatively large volumes of particles of uniform size and shape is independent of the size of the particles, provided surface effects are inconsiderable. When the specific surface area becomes large, porosity tends to increase with reduction of particle size. In particulate bulks comprising several sizes, the presence of isolated large particles tends to reduce the total porosity of the bulk and a similar effect results from the introduction of particles small enough to occupy voids without necessarily disturbing the structure. The situations arising from the mixing of different sizes of particle are so complex in real, as opposed to special and hypothetical, circumstances that pursuit of the theoretical argument is hardly justified by its fruits.

It is known in a general sense that the porosity of a poured mass is reduced by increasing the height from which the particles are dropped into the receiver and by reducing the density of the descending cloud of particles. Equally, it has long been appreciated that the voidage in the immediate neighbourhood of a container wall is greater than that of the whole assemblage of particles. Available experimentation in this field is, however, inconclusive.

The extensive literature upon which these outline comments are based is well represented in the contributions noted in the References. From these and others it is apparent that the major variables which influence the packing of dry granular solids are those named in Table 1. The table shows the authors' assessment of the type and weight of the evidence at present available. Typical sources are also indicated by numbered references.

It is evident that although theoretical considerations have indicated the possible significance in packing of the properties of the materials of particles and containers, experimentalists have tended to ignore these features. This paper records an attempt to explore, by controlled and defined experiment, the independent effects of the more important variables of deposition upon packing, and the modifying influence of the intrinsic properties of the particle material upon the trends produced. The ultimate objective has been the statement quantitatively of an explicit general mechanism of packing. The intrinsic properties considered are (i) surface friction at low levels of kinetic energy of the particles, and (ii) resilience, regarded as likely to be of increasing importance as the kinetic energy of the constituent particles of a bed increases.

* Now at Shell Research Limited, Thornton, Cheshire.

Table 1. *Evidence in the literature concerning variables considered to be important in the packing of particles*

	Experimental		Theoretical	
	Evidence	Ref. No.	Evidence	Ref. No.
1. *Particle*				
Shape	**	3, 4, 6	**	3, 6
Absolute size	*	7, 14	*	6
Size distribution	**	1, 6, 7	**	6, 7, 15
Mass	None		*	8
Surface friction	None		*	6
Elasticity	None		*	8
2. *Container*				
Shape	*	5, 8	*	8
Size	**	2, 3, 11	*	8, 13
Surface friction	None		None	
Elasticity	None		None	
3. *Mode of deposition*				
Intensity	**	9	*	8
Method	*	9	*	8
4. *Energy input to the mass*				
During deposition	**	4, 9	*	8
After deposition	**	10, 12	*	10, 12

1, Anderegg 1931; 2, Brown and Hawksley 1946; 3, Carman 1937, 1938; 4, Coulson 1949; 5, Denton 1957; 6, Fraser 1935; 7, Furnas 1929, 1931; 8, Graton and Fraser 1935; 9, Kolbuszewski 1950; 10, Macrae, Finlayson and Gray 1957; 11, Rose 1945; 12, Stewart 1951; 13, Verman and Banerjee 1946; 14, Westmann and Hugill 1930; 15, White and Walton 1937.

* qualitative evidence of the effect of the variable is available.
** some quantitative evidence of the effect of the variable is available.

2. Method

2.1. *Materials studied*

In order to minimize the shape factor, spherical particles were used. The significance of shape was disclosed by using, in one case, two batches of particles of slightly differing sphericity. In all cases the particles were of nearly uniform size. The materials used were steel ball bearings, phosphor-bronze ball bearings, lead shot, polystyrene beads and two forms of glass beads. Glass beads A were more nearly spherical than glass beads B. The size and shape characteristics of the six materials are shown in Table 2, together

Table 2. *Particle size, shape, density and friction*

Material	*Size (in.)*	*Range of size (in.)*	*Axial ratio*	*Density (g/cm³)*	*Friction index*
Steel	0·125	±0·0002	1·00	7·78	1·15
Phosphor-bronze	0·125	±0·0002	1·00	8·92	1·19
Lead shot	0·119	±0·008	1·05	11·27	1·21
Glass A	0·125	±0·009	1·04	2·524	1·13
Glass B	0·118	±0·010	1·10	2·529	1·30
Polystyrene	0·121	±0·009	1·08	1·023	1·39

with their densities and an index of the surface friction. The last feature was assessed by an empirical measurement involving both rolling and sliding components. The particles were allowed to flow through a circular orifice (0·540 in. diameter) in the base of a 3 in. cylindrical container. The rate of flow was expressed in time per unit volume. This empirical measurement is doubtless influenced by variations in the shape of the particles and is most truly a measure of surface friction when particles are most perfectly spherical.

The resiliences of all the materials except lead, expressed as coefficients of restitution and regarded as a measure of their ability to impart kinetic energy to a target on impact, are shown in Fig. 1. The material used as anvil, or striking

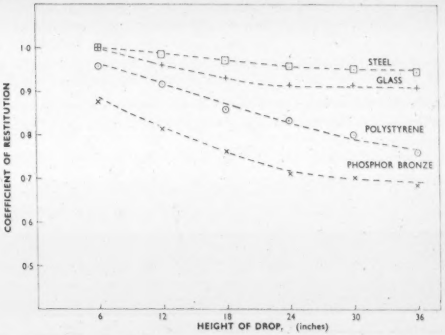


Fig. 1. Coefficients of restitution of materials studied.

plate, in determining these values was flint glass whose coefficient of restitution (flint glass on flint glass) is very nearly unity.

2.2. *Packing method*

The better to ensure understanding of the events, a carefully simplified regime was used whereby beds, some 45 cm high, were built by pouring the particles at a controlled rate into a 3 in. diameter vertical glass tube with a horizontal base plate of stainless steel. The cylindrical hopper moved inside the glass receiver to allow small heights of drop to be used and a cylindrical delivery tube, 1 in. in diameter, extending from the controlled hopper outlet to a distance of 1 in. above the growing bed, ensured that particles reached the surface of the growing bed in a consistent manner at all heights of drop, and that the descending particles impinged upon a substantially constant area. The intensities of deposition named in the succeeding text refer to the number of particles which fell per second upon this constant area, whose extent is difficult to define precisely but is approximately 1 in^2. Because of this slight uncertainty the rates of deposition are not stated in terms of area. To keep the height of fall of the particles at a constant predetermined value throughout an experiment, the hopper could be moved upward or the receiver downward at the same rate as the surface of the growing bed rose. The closing mechanism in the base of the hopper was manually controlled from the top and allowed intensities of deposition in the range 450 to 9600 particles per second. Very low intensities of deposition were achieved by using a subsidiary mechanical feed in the form of a toothed wheel which, rotating at selected speeds, delivered particles individually to the main hopper at rates down to seven particles per second.

2.3. *Packing criteria*

Three criteria of packing were used, namely the overall porosity of the whole bed, the porosity and position of incremental volumes throughout a bed and the spatial disposition of individual particles as disclosed by transverse

sections. The use of contact point counts and their interpretation as co-ordination numbers was considered and rejected, as was the possible description of unit voids.

2.3.1. *Contact points, co-ordination numbers and unit voids.* The use of a marking technique, similar to those of Smith, Foote and Busang (1929) and Bennett and Brown (1940), to indicate by chemically produced marks on the surface of each sphere the immediate proximity of adjacent spheres, involves the interpretation of these marks as co-ordination numbers and has in the past been used as a measure of porosity. This is unwarranted and can be misleading to a major extent. While it is true that in the various homogeneous, systematic assemblages of spheres studied theoretically by, for example, Graton and Fraser (1935) all the spheres have the same co-ordination number, these packings are not defined by the co-ordination number alone but only by that in conjunction with the shape of the unit cell. A packing of ideal spheres in three dimensions can be distorted without alteration to the co-ordination number, but with significant change in the porosity. For example, if a sphere is touched by five others the packing can be distorted without change in this co-ordination number so that the porosity changes in the range 59·7% to 47·7%.

The interpretation of an experimentally determined co-ordination number is further complicated by experimental imperfections, in that real particles are never ideally shaped and sized. In consequence, one situation of several which may arise is that in which a sphere, slightly smaller than its neighbours, may rest on three others and show three contact points, while adjacent spheres which just fail to touch the small one show more than three contacts. None shows the number which truly is associated with the packing.

In the case of spheres which have been poured into a container the unit void may be any one of an infinite number of polyhedral forms. Consequently, it appears that the definition and assessment of the individual voids (see Lockwood 1950, Nuss and Whiting 1947) would yield data of such detail and complexity that more important overall trends would be obscured.

2.3.2. *Direct measurement of porosity.* The measurements of overall porosity were made both (i) by measuring the voidage by means of a displacement liquid (a very dilute aqueous solution of a detergent) inserted through the base of the receiver, and (ii) by measuring the height and weight of the bed, the area of the receiver and the density of the particle material. At the 40% level the accuracy of a porosity measurement was to $\pm 0{\cdot}1\%$. Characteristic levels of porosity are produced by different materials in response to specific conditions of deposition. For any material these levels lie within a range which is peculiar to the material. Consequently the overall porosity, expressed as a proportion of void in the bulk, cannot of itself allow valid comparisons of the response of different materials to defined environmental conditions. Instead, we have used for such comparisons the percentage improvement in packing produced by specified conditions, the datum for any material being the porosity of a bed of that material produced solely by cascading* conditions of deposition. This is derived from the suggestion by one of us (Macrae 1958) that only two situations of the packing of a mass of particles can be explicitly stated, namely (i) that in which the mass is so unstable that any small decrease of stability would cause the formation of a cloud, and (ii) that in which the particles are close packed. Any other state of packing refers to a situation of intermediate stability uniquely determined by the mechanical treatment experienced by the mass. The state of the cascaded mass is a practicable and tolerably reproducible situation close to the former definable state. The improvement in packing is

$$\frac{(V_p/V_b)_d - (V_p/V_b)_c}{1 - (V_p/V_b)_c}$$

where V_p is the volume of the particles, V_b is the volume of the bed, c refers to cascading conditions of deposition, d refers to another, specified condition of deposition.

The porosities of successive layers of a bed were measured, initially, by the use of a displacement aqueous solution of low surface tension.

2.3.3. *Evaluation of sections.* In order to explore the variations in porosity in successive annular incremental volumes, or layers, and to establish directly the kind, magnitude and extent of the disturbances of the packing due to the existence of a wall, horizontal layers of beds were fixed by the introduction, as an impregnating medium, of Araldite 900, a low-viscosity liquid casting resin which sets at room temperature. Sections of this body having been cut, trends in the packing of spherical particles from a selected plane were explored by measuring the distance of each particle centre from the reference plane. These data were then grouped to reveal the trend in the packing. In random packings the choice of interval between successive, parallel cursor planes appears to be arbitrary and in these experiments d, the particle diameter, was used. With an ordered, or near ordered, array of spherical particles the interval d is not appropriate. In these cases the interval, found by trial, which allowed only one layer of particles to be grouped in each interval was used. The two operations of measuring the distance of each particle centre from the reference plane and of grouping the data were condensed into one operation of counting by the following method. The image of a series of concentric circles, spaced as dictated by the circumstances noted above, was focused upon a section and the number of particles cut by each circle counted. The number of particles cut by unit length of the line was used as a basis for comparison of the porosities of successive cylindrical layers. The difference in the porosities of the entire bed held by the container and of that part not affected by the wall was estimated in terms of porosity thus:

a = number of particles counted in a total section of diameter 3 in.

b = number of particles counted in a core of the section of diameter less than 3 in.

$$\frac{a}{\text{total area of section}} \propto (1 - \text{overall porosity}), \text{ which is known}$$

$$\frac{b}{\text{area of core}} \propto (1 - \text{porosity of core}), \text{ which is to be found.}$$

The standard error of the porosity difference between a core and the whole bed calculated from the sample variations (ten sections) was 0·4%.

3. Results and comment

3.1. *Bed and layer porosities*

3.1.1. *Cascaded beds and surface friction.* The datum values of the porosities of beds formed by cascading are

* *Cascading*: in these experiments, a movement of particles by sliding and rolling without free fall.

shown in Table 3, together with the indices of friction of the particles and their axial ratios.

Table 3. *Porosity of cascaded beds and surface friction of materials*

Material	*Porosity (%)*	*Friction index*	*Axial ratio*
Polystyrene beads	40·9	1·39	1·08
Phosphor-bronze	40·9	1·20	1·00
Steel	40·4	1·15	1·00
Glass A	40·1	1·13	1·04
Glass B	40·1	1·30	1·10
Lead	39·3	1·21	1·05

There is here some suggestion that friction plays an experimentally significant part in packing processes under cascading conditions, but no precise comparison can be made between packing systems of particles possessing different degrees of sphericity. Simple surface friction effects are likely to be obscured by even small resilience effects when particles are deposited by free fall.

3.1.2. *Layer porosities.* Little variation could be found in the porosity in successive 2 in. thick layers of beds of glass, phosphor-bronze and lead particles cascaded into the receiver or dropped from heights up to 18 in., the basal layer being ignored. As the thickness of layers was reduced, so the range of variation increased from values of 0·1% porosity between all the 2 in. layers of a bed, to 0·25% between all the $\frac{1}{8}$ in. layers. These variations indicate the propriety of measuring the overall porosity as an indication of the packing, and suggest that, in a general sense, packing is homogeneous throughout the height of a bed contained in a cylinder. This comment assumes a consistent wall effect in regions remote from the base. The selection of very small incremental volumes in the form of horizontal layers simply displays the infinitely detailed character of the voidage and does not reveal any general trend related to the position of an increment in the bed.

3.1.3. *Overall effects of depositional variables.* Figs 2, 3 and 4 show the consequences of varying the height from

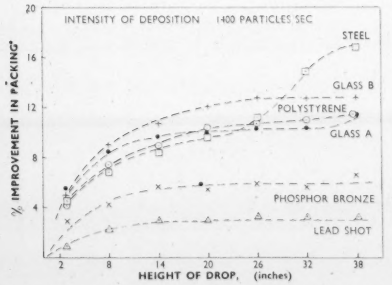


Fig. 2. Variation of improvement in packing with height of drop at 1400 particles per second.

which particles of steel, phosphor-bronze, lead, glass and (Fig. 2) polystyrene were dropped into the receiver and of varying, also, the intensity of deposition. Since resilience is effective when particles collide, increasing the momentum of the descending particles must result in an increased resilience effect. It will later be shown that these figures contain important evidence which leads to an hypothesis of the packing mechanism. For the moment the following observations are pertinent:

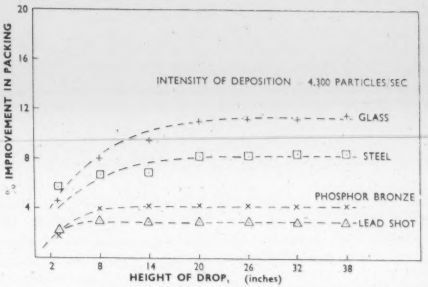


Fig. 3. Variation of improvement in packing with height of drop at 4300 particles per second.

(i) With all the materials and over the range of heights of drop used, the overall porosity decreases with increasing height of drop at constant intensity of deposition. Except in the case of steel dropped at 1400 particles per second, an equilibrium value of porosity occurs at a critical height above which no further decrease of porosity occurs.

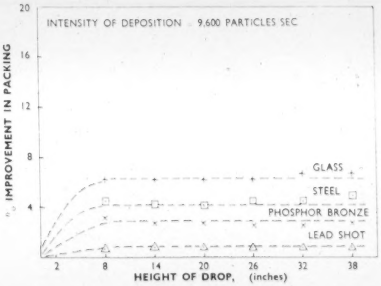


Fig. 4. Variation of improvement in packing with height of drop at 9600 particles per second.

(ii) This trend is caused by increasing the kinetic energy of the particles at impact. It is accentuated in the lowest intensity of deposition shown (Fig. 2) and the improvement in packing is reduced by increasing the intensity of deposition.

(iii) The steel spheres, at the lowest intensity shown (1400 particles per second), pack more closely than the other materials. The trend towards closest packing in this case is strikingly different from the trends shown by the other materials and by steel spheres at higher intensities of deposition. A possible exception to this statement is the behaviour of the glass beads A (more nearly spherical than glass beads B) which appear, at a height of drop of 36 in. and a deposition rate of 1400 particles per second, to be incipiently following the trend of the steel spheres. Visual observation showed that at heights of drop exceeding 18 in. the steel spheres tended to pack, not only closely, but also in an ordered fashion, at any rate in the visible outer layers of side and top. Further experiments with glass beads A, glass beads B (two highly resilient materials of different sphericity) and phosphor-bronze spheres (almost perfectly spherical but of intermediate resilience) confirmed and amplified the observation above concerning glass beads A.

These packed in orderly array at the wall and top when dropped from 90 in. at 1400 particles per second. The less spherical glass beads B did not pack in orderly array at heights of drop up to 126 in. and phosphor-bronze failed to do so at heights up to 25 ft, although a small improvement in packing was apparent. The data quantifying these trends are in Table 4.

Table 4. *Improvement in packing of glass beads A and phosphor-bronze at heights of drop between* 36 *in. and* 300 *in.*

Height (in.)	36	42	54	60	66	90	126	180	300
	Improvement in packing (%)								
Glass beads A	11·0	11·2	11·5	14·0	14·0	14·0	—	—	—
Phosphor-bronze	6·1	—	—	—	—	8·1	7·8	8·3	8·1

It is evident that both the general trend towards closer packing with increased height of drop and the importantly anomalous behaviour of steel spheres and glass beads A in

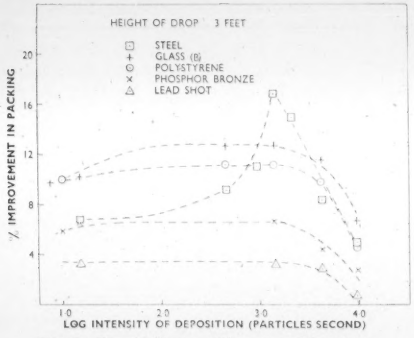


Fig. 5. The significance of intensity of deposition.

achieving ordered packing are related to the intensity of deposition. The significance of this depositional variable is further shown in Fig. 5. The known trend that the porosity of a bed is reduced as the intensity of deposition is reduced (Kolbusczewski 1950) is seen here to be only a partial statement of the truth. In fact, there is displayed for each material, except lead shot, a critical range or value of intensity of deposition which produces closest packing at the prevailing height of drop. Intensities not only above but also below this range or value result in the production of more open structured beds. The apparently anomalous behaviour of lead shot is evidently no more than an extreme situation of the behaviour of the other materials.

3.2. *Wall effects in a cylinder*

3.2.1. *Distribution of porosity from the wall.* All the above values of porosity refer to the entire bed as it is held in the receiver. Such a bed can be regarded as comprising essentially two main parts. These are a core of 'normal' packing which has the structure of an infinitely extensive bed, and an annulus of disturbed packing whose structure is associated with the wall. To give the data now available a significance wider than merely the environment of a 3 in. diameter glass tube, and to justify a generally applicable quantitative statement of the significance of the resilience of the particles, it is necessary at this stage to measure the effect upon the packing of the presence of the wall.

The packing situations studied experimentally by the examination of sections of beds were those produced by lead shot, phosphor-bronze spheres and polystyrene beads when either cascaded or dropped from 36 in., and glass beads A when either cascaded or dropped from 36 in. or 90 in. The last situation, consisting of ordered packing, was studied in place of that produced by steel spheres dropped from 36 in. because of our inability to cut the sections through this hard material without disrupting the specimen. The porosities ascribed to successive layers from the wall in each case are set out in Table 6. This table may be regarded as a form of histogram, the vertical lines indicating distance from the wall being on a scale of distance.

The calculated wall effects shown in Table 6 for the ordered packing of steel spheres were derived from observations of a bonded, but not sectioned, bed and from the

Table 6. *Porosities of successive annuli, or layers, showing the measured wall effects with four materials*

Material	*Deposition*	*Overall porosity*	*Distance from wall (in.)* 0	1/4	1/2	3/4	1				
Lead	cascaded	39·3	47·3	40·3	36·6	37·5	37·0	38·9	37·3	37·7	
	36″ drop	38·0	44·0	35·3	34·2	37·7	38·4	36·4	36·6	38·7	
Phosphor-bronze	cascaded	40·9	49·0	39·3	37·8	38·1	40·1	39·3	40·2	40·3	
	36″ drop	38·2	45·7	37·7	33·1	34·0	38·3	37·4	37·3	37·9	
Polystyrene	cascaded	40·9	47·8	38·8	39·2	38·7	40·3	41·1	40·6	40·8	
	36″ drop	36·3	45·0	34·0	29·1	33·8	37·0	37·1	37·5	35·5	
Glass A	cascaded	40·1	47·9	39·3	39·0	36·4	37·8	38·8	38·5	37·3	
	36″ drop	35·6	42·4	37·9	29·8	31·5	34·5	36·1	34·9	36·6	
	*90″ drop (ordered)	35·0	37·7	33·2	33·2	33·2	31·4	34·4	33·9	37·6	37·8
Steel†	36″ drop	33·6	38·0	32·0							36·3

* In this case, the ordered packing, note that the interval or thickness of successive layers or slices was not d as used in the other cases. † Calculated values.

measured wall effect data for the ordered packing of glass beads A thus:

	Porosity
Outer layer	38·0 (data for glass)
7 succeeding layers	x (unsectioned steel)
Core	36·3 (data for glass)
Overall	33·6 (data for steel)

These data show that with each material cascading produces a packing of high porosity at the wall, a lower porosity in the second annular layer and a nearly constant value throughout the remainder of the bed. In cascading the particles roll to the wall which does not interact with them to exert a far reaching effect upon the packing. The data, however, show that the overall porosity is comparatively high because of the large voidage at the wall. When dropped, the particles tend to pack from the wall. Because in these experiments the cylindrical wall approximates in shape to one side of a unit cell, better packing is promoted. The improved packing at the boundary allows closer packing in a few of the succeeding layers. The influence of the wall is no longer apparent in four layers and the packing has again become that of an infinitely extensive bed.

A different situation is evident when all the circumstances allow ordered packing at the wall. As the data show, this ordered packing in the first layer allows a similar packing in succeeding layers until, eventually, disorder developes in a small central core. Doubtless the last arises from the interaction and multiplication of dislocations emanating from the wall and from imperfections of the receiver in relation to the shape and size of the unit cell. It is proper to emphasize an important distinction here. The disordered packing in the core of an otherwise ordered bed is not as compact as optimum, but haphazard, packing. The authors suggest that the former disordered core is not that which would occur in an infinitely extensive bed, but is produced uniquely as a consequence of the order in the packing of the outer layers, and in a larger container would still be confined to a core of substantially the same size as that revealed in the 3 in. container. Consequently, the ordered array would probably be the extensive, characteristic, packing of such a bed. A bed built inwards from an ordered packing at the wall would thus comprise three regions, namely a single layer of high porosity at the wall, a small (around 1 in. diameter) core of fairly high porosity and disordered packing and, between these, an annular ordered bed of lower porosity, filling the remainder of the receiver irrespective of its size.

3.2.2. *Porosity of the infinitely extensive bed.* If the hypothesis be accepted that in the absence of ordered disposition at the wall the packing in a cylindrical receiver can be regarded as an annulus of disturbed packing, determined by the wall, surrounding a region of undisturbed, 'normal' packing which is that which would have been established in an infinitely extensive bed, then the generally applicable quantitative improvement in the normal packing caused by the selected conditions of deposition can be assessed. The differences between normal and container-held (that is, entire) packings produced by cascading are given in Table 7, and those for packings produced by dropping the particles from a height of 36 in. in Table 8.

The comparatively small differences between the porosity of the normal (i.e. infinitely extensive) packing and that of the entire container-held bed shown in Table 8 result from the self-cancelling tendency of two opposing trends, which occur when a bed is created by dropping the particles from a height. In such circumstances, it is suggested, the wall acts (i) as a starting surface tending to promote good packing of rebounding particles, and (ii) as a limit or boundary at which, because of shape and size, particles cannot pack so closely as they can against other particles.

Table 7. *Porosities of normal packings, N, and container-held beds, c, produced by cascading*

	Porosity (%)		
Material	*c*	*N*	*c–N*
Lead shot	39·3	37·5	+1·8
Phosphor-bronze	40·9	39·7	+1·2
Glass beads A	40·1	38·3	+1·8
Polystyrene beads	40·9	39·8	+1·1

Table 8. *Porosities of normal packings, N, and container-held beds, c, produced by dropping particles 36 in.*

	Porosity (%)		
Material	*c*	*N*	*c–N*
Lead shot	38·0	37·7	+0·3
Phosphor-bronze	38·2	37·6	+0·6
Glass beads A	35·6	35·5	+0·1
Polystyrene beads	36·4	36·7	−0·3

All the porosities of the normal cascaded packings are approximately 1·5% lower than those of the container-held beds. Normal packings produced by dropping the particles from 36 in. are 0·2% lower in porosity than the container-held beds. Consequently the improvement of a normal packing caused by increasing the impact velocity of falling particles to 150 in/s is approximately 1·3% of porosity lower than the improvement noted for the container-held beds. That is, filling a receiver by dropping particles into it results in a greater improvement of the packing at the wall than at the core, compared with filling by a gentle cascading operation.

3.3. *The influence of resilience*

It is now possible to make a first quantitative and fundamental appraisal of the influence of the resilience of the particle material upon the packing. The optimum (i.e. closest) packing of steel spheres, glass beads, phosphor-bronze spheres, lead shot and polystyrene beads is related to the coefficients of restitution of the materials in Fig. 6. The relationship is both in terms of porosity and of improvement in packing. This optimum packing is that which would be achieved for all particle materials at a level of intensity of deposition in the range 100 to 1400 particles per second and at impact velocities greater than 150 in/s. The optimum packing may either be that of the extensive core free from wall effects in the case of haphazard packing, or that of the extensive ordered packing between the single boundary layer and the small disordered core, a situation achieved with steel and glass A at 1400 particles per second.

It is evident from Fig. 6 that the resilience of the particle material must be regarded as having as much influence upon the porosity of a bed as the conditions of deposition.

4. A mechanism of packing

A fuller exploration of the role of the resilience of the particle material stems from the visual observation that the activity of steel spheres at the surface of a growing bed packing in ordered array exceeded that of phosphor-bronze spheres dropped from the same height. The latter packed in haphazard manner. A slow motion (64 frames per second) cine film record of the growth of beds of steel spheres, built by dropping the particles from heights of 18 in. and 36 in., and of a bed of phosphor-bronze spheres, built by dropping

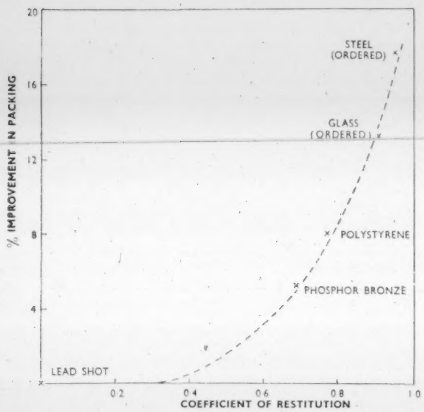


Fig. 6. The relationship of optimum packing to resilience.

the particles from 36 in., all at an intensity of 1400 particles per second showed that four regions of the growing bed could always be distinguished:

(*a*) a stable region in which all particles are at rest;
(*b*) a region comprising the top layers of the bed in which particles move in groups or aggregates;
(*c*) a dome-shaped region above the last in which all the particles are in individual motion, cascading or interacting;
(*d*) above the dome, a cloud of rebounding and arriving particles.

Measurements of the extent of the last three regions and visual assessments of the activity in them are summarized below.

Phosphor-bronze dropped from 36 *in.*

Region (*b*)—At least six layers thick. Motion sluggish and intermittent.
Region (*c*)—$\frac{1}{4}$ in. high. Motion mostly cascading.
Region (*d*)—$\frac{3}{4}$ in. high. Sparsely populated.

Steel dropped from 18 *in.*

Region (*b*)—At least five or six layers thick. Motion continuous.
Region (*c*)—$\frac{1}{2}$ in. high. Activity greater than phosphor-bronze dropped from 36 in.
Region (*d*)—$1\frac{1}{4}$ in. high. Sparsely populated.

Steel dropped from 36 *in. and achieving ordered packing*

Region (*b*)—At least six layers thick. Motion continuous and notably more vigorous than in the two preceding cases.
Region (*c*)—More than $\frac{1}{2}$ in. high. This region is now in fact a very densely populated cloud, i.e. rebounding dominates cascading.
Region (*d*)—$1\frac{1}{4}$ in. high. Much more populated than in the previous experiments.

The most striking difference in the behaviour of the particles in the three situations examined is that of the intensity of the jostling of the particles in the top layers of the growing bed. This was greatest for steel spheres dropped from 36 in. and it was evident from the film that this movement, having a critical intensity, caused these particles to pack in preferred positions. It is reasonable to conclude for the moment that any material will achieve optimum packing if a critical level of activity in the upper layers of a growing bed is exceeded. Because materials differ in physical properties, the extent to which kinetic energy, originating in the energy of fall, is transmitted through a bed depends upon the material of which the particles are composed. Consequently, for each material there is a critical level of energy increment (energy increment being the kinetic energy imparted to a target particle by a descending particle) which must be exceeded to produce the required intensity of activity. Further, it is clear that a critical rate of repeated applications of the appropriate level of energy increment must be exceeded if the required intensity of continuous movement in the upper layers is to be achieved. For a particular intensity of deposition, the rate of repeated applications of energy increments depends upon the elastic properties of the particle material.

The intensity of activity in the upper layers of a growing bed may be assessed as the time during which an energy increment is effective in maintaining movement in the *n*th layer below the surface. Thus the condition for optimum packing can be stated as a requirement to exceed a minimum magnitude of motion for a minimum time. It follows directly that the minimum time may not be achieved (i) if the energy increments are less than the critical, (ii) if, at an energy level exceeding the critical, the rate of renewed application of increments is reduced below a critical value, (iii) if, at energy increment and rate of repeated application levels exceeding the critical, a critical upper level of intensity of deposition is exceeded.

The last restriction follows from the fact that increasing the rate of deposition increases the number of target particles as well as the number falling, so that the number of impacts per second per target particle tends to remain constant while the rate of growth of the bed increases in the proportion of the increase in the rate of deposition. This results in a reduction in the duration of upper layer activity with the consequent increase in porosity of the bed at high deposition rates noted in Fig. 5. For any material the minimum time of motion, achieved by appropriate height of drop and intensity of deposition, depends upon the ability of the particle material to impart kinetic energy on impact, that is, upon its resilience.

While these conditions determine optimum packing and the reasoning accounts for the general trends displayed in Figs 2, 3, 4 and 5, the production of ordered packing as displayed, for example, by the steel spheres in particular circumstances involves another condition, namely the existence of a wall interacting with sufficiently active particles. Thus an improvement in packing can occur if the above conditions are met either (i) by the interaction of particles only or (ii) by the interaction of particles with other particles and with the receiver walls. The former produces a haphazard packing which tends towards an optimum closeness and is that of an infinitely extensive bed. The latter tends to packings which originate from the wall and tend to be ordered. Minor defects in particle shape inevitably prevent the creation of ordered packings in any circumstances, a consequence of the hypothetical mechanism illustrated by the behaviour of the nearly spherical glass beads A and the less spherical glass beads B.

The authors' attempt to express quantitatively the conditions which result in optimum packing assumes that the hypothetical mechanism outlined above is correct and that the rate of decay of the kinetic energy imparted by an

impacting particle is a function of resilience only. The value of this decay rate depends upon the number of collisions and the kinetic energy lost in each. The general statement is $V_x = U_m\{f(e)\}^n$, where V_x is the velocity of particles in the xth layer, U is the velocity of the impacting particle, m refers to the particle material, $f(e)$ is a function of the coefficient of restitution, n is the number of collisons. If it be assumed that collisons are all simple, the general expression becomes $V_{nm} = U_m\{(1 + e)/2\}^n$, where V_{nm} is the velocity of the particle in the nth layer, i.e. $n = x$. (It is virtually certain that n is greater than x.) Assuming for the moment that $n = x = 6$ (least number of layers observed in movement when steel spheres packed in ordered array after falling 36 in.), the impact velocity at which perfect glass spheres and perfect phosphor-bronze spheres may be expected to pack in ordered array can be calculated from the data for steel: thus, $U_s r_s^6 = U_g r_g^6$, where $r = \frac{1}{2}(1 + e)$ and g and s refer to glass and steel. The respective impact velocities and required heights of drop are

glass, 15·8 ft/s, i.e. 47 in. drop
phosphor-bronze, 56 ft/s, i.e. 50 ft drop.

In experiments the glass beads A achieved ordered packing when dropped from 90 in., the less truly spherical glass beads B did not show ordered packing even when dropped from 125 in. and the phosphor-bronze spheres did not pack in this manner when dropped from 25 ft, the greatest height available to us. Clearly, x is greater than 6. An estimate of the value of x derived from the data for glass beads A is $x = 12$.

The general statement may now be made on the basis of our data and argument that ordered packing of a material m of resilience e_m will be induced by an impact velocity U_m, when $U_m = 120/\{\frac{1}{2}(1 + e_m)\}^{12}$ in/s and will be prevented by shape imperfections.

5. Conclusions

5.1. *Wall effects*

The effect of a wall upon the packing in its vicinity is critically dependent upon the conditions of deposition. When particles are cascaded, the packing proceeds from the centre towards the wall whose perturbing effect extends only two particle diameters. The porosity of the extensive core is less than that of the whole bed. In a 3 in. diameter tube the amount of this difference is 1·5% porosity. Beds formed when particles fall from heights greater than 30 in. may consist, like cascaded beds, of a normal core surrounded by an annulus in which the wall effect persists for four particle diameters. In such beds the porosity of the core is only slightly lower than that of the whole bed. In a 3 in. diameter tube this small difference has a value of 0·2% porosity. Sufficiently intense interaction of moving particles in the upper layers of a growing bed with one another and with the wall allows ordered packing at the wall. The bed packs in this case from the wall inwards in ordered array, until minor imperfections create a small disordered core.

5.2. *Packing of an infinitely extensive bed of spheres*

When formed by cascading, the porosity of an infinitely extensive bed of dry spheres is in the range $38{\cdot}5 \pm 1{\cdot}0\%$. The packing depends mainly upon the spherical shape and to only a minor extent upon the surface friction of the particles. The extent of compaction of a bed of spheres from an initial porosity level within this range is equally determined, *inter alia*, by the height of drop, intensity of deposition and resilience of the particle material. Minor departures from perfect sphericity mask the modifying influence of resilience upon the packing trends produced by the conditions of deposition.

5.2.1. *Conditions of deposition.* Optimum packing of an infinitely extensive bed of spherical particles of all materials is achieved above a height of drop of approximately 30 in. (impact velocity 150 in/s) at an intensity of deposition in the range 100 to 1400 particles per second.

5.2.2. *Resilience.* The level of optimum packing which is produced by the conditions of deposition depends upon the resilience of the particle material. The salient features of the relationship between the resilience, as measured by the coefficient of restitution, and the improvement in packing with reference to the core of a cascaded bed as datum are:

Coefficient of restitution	0–0·4	0·4–0·8	0·8–1·0
Improvement in packing (%)	Nil	nil–10	10–*c.* 20

5.3. *The activation of ordered packing*

For an optimum packing the energy increment supplied by a falling particle on deposition must be above a critical value, the rate of renewed application of increments must be above a critical level and the intensity of deposition must be below a critical level.

The presence of a wall interacting with the moving upper layers of a growing bed allows ordered packing at the wall if the interaction is sufficiently great. Ordered packing at the wall promotes ordered packing in the succeeding, inner layers. In the circumstances studied, ordered packing occurs when the intensity of deposition is approximately 1400 particles per second at impact velocities above a critical value. This value has tentatively been related to the resilience of the particle material in the expression

$$U_m = \frac{120}{\{\frac{1}{2}(1 + e_m)\}^{12}} \text{ in/s.}$$

Ordered packing is prevented by minor imperfections of the particle shape.

References

ANDEREGG, F. O., 1931, *Industr. Engng Chem.*, **23,** 1058.
BENNETT, J. G., and BROWN, R. L., 1940, *J. Inst. Fuel*, **13,** 232.
BROWN, R. L., and HAWKSLEY, P. G. W., 1946, *Nature, (London)*, **157,** 585.
CARMAN, P. C., 1937, *Trans Instn Chem. Engrs*, **15,** 150.
—— 1938, *Trans Instn Chem. Engrs*, **16,** 168.
COULSON, J. M., 1949, *Trans Instn Chem. Engrs*, **27,** 237.
DALLAVALLE, J. M., 1948, *Micromeritics* (New York: Pitmans).
DENTON, W. H., 1957, *The Packing and Flow of Spheres*, E/R 1095 (Atomic Energy Research Establishment).
FRASER, H. J., 1935, *J. Geology*, **43,** 910.
FURNAS, C. C., 1929, *U.S. Bureau of Mines Bull.*, No. 307.
—— 1931, *Industr. Engng Chem.*, **23,** 1052.
GOETZEL, C. L., 1949, *Treatise on Powder Metallurgy* (London: Interscience Publishers Limited).
GRATON, L. C., and FRASER, H. J., 1935, *J. Geology*, **43,** 785.
HEYWOOD, H., 1946, *J. Imp. Coll. Chem. Engng Soc.*, **2,** 9.
KOLBUSCZEWSKI, J., 1950, *Research*, **3,** 478.
LOCKWOOD, W. N., 1950, *Bull. Amer. Ass. Petrol. Geol.*, **34,** 2061.
MACRAE, J. C., 1958, *Conference of Science in the Use of Coal (Sheffield)*, B19 (Institute of Fuel).

MACRAE, J. C., FINLAYSON, P. C., and GRAY, W. A., 1957, *Nature* (*London*), **179**, 1365.
MELDAU, R., and STACH, E., 1934, *J. Inst. Fuel*, **7**, 336.
NUSS, W. F., and WHITING, R. L., 1947, *Bull. Amer. Ass. Petrol. Geol.*, **31**, 2044.
ROSE, H. E., 1945, *Proc. Instn Mech. Engrs*, **153**, 141.
SMITH, W. O., FOOTE, P. D., and BUSANG, P. F., 1929, *Phys. Rev.*, **34**, 1272.
STEWART, D. A., 1951, *High Quality Concrete* (London: Spon).
VERMAN, L. C., and BANERJEE, S., 1946, *Nature* (*London*), **157**, 584.
WESTMANN, A. E. R., and HUGILL, H. R., 1930, *J. Amer. Ceram. Soc.*, **13**, 767.
WHITE, H. E., and WALTON, S. F., 1937, *J. Amer. Ceram. Soc.*, **20**, 155.

Porosity and surface area of a granular bed from measurements of the flow of air through the bed and measurements of the granular shape factors

by W. M. JONES, D.Phil., F.Inst.P., and D. MILLS, B.Sc., Grad.Inst.P., Department of Physics, University College of Wales, Aberystwyth

MS. received 21*st December* 1959, *in revised form* 13*th October* 1960

Abstract

Flow experiments give an equation relating ϵ *to* S *in terms of measurable quantities, where* ϵ *is the inter-granular porosity of the bed, and* S *is the envelope surface area of the granules in* 1 cm^3 *of bed. Heywood's shape factors for the granules give a second equation relating* ϵ *to* S. *Using the two equations* ϵ *and* S *can be determined uniquely. If the tortuosity factor be taken to be* $\pi/2$ *in the equations given previously* (*Jones* 1956), *connecting the friction factor* λ *and the Reynolds number* (*Re*), *then* $\lambda = 16/(Re)$ *for granular beds, as for long straight tubes.*

THE flow of a fluid through a granular bed is related to ϵ^3/S and other measurable quantities, where ϵ is the inter-granular porosity of the bed, and S is the envelope surface area of the granules in 1 cm^3 of bed. ϵ and S cannot be determined uniquely from flow alone and when the granules are themselves porous difficulties exist in determining either ϵ or S separately by other means; for example, if ϵ be determined by a liquid displacement method then doubt exists as to the extent of the penetration of liquid into the pores of the granules and if attempts are made to seal the pores before immersion in liquid doubts exist as to the efficiency of the sealing. S cannot be determined by an adsorption method because then the internal area of the pores would contribute to the area which is measured. However, Heywood's (1947) shape factors can be used to give a set of equations connecting ϵ with S and these can be used in conjunction with flow experiments, to determine ϵ and S uniquely.

Theoretical

The principle of the method is the application of the result of the previous paper (Jones 1956), namely, that if the flow resistance of a bed is represented on a graph showing $\log \lambda$ as a function of $\log (Re)$, where λ is the friction factor and (Re) the Reynolds number, then the graphs will be the same for beds of irregular granules as they are for beds of regular granules. To represent the flow resistance on such a graph use is made of the equations

$$\lambda = \sqrt{2}\Delta p \epsilon^3 g l/\rho u^2\{S + (4/D)\}$$
$$\text{and} \qquad (Re) = 2\sqrt{2}u\rho/\eta\{S + (4/D)\} \qquad (1)$$

where Δp is the flow resistance of the bed (cm w.g.), ϵ is the porosity of the bed, D its diameter, and l its length, S is the envelope surface area of the granules in 1 cm^3 of bed, u is the velocity of flow of the fluid of density ρ and viscosity η into, or out of, the bed. For beds of regular solid particles (e.g. lead shot) all the quantities in Eqn (1) can be determined and the position of the graph established. The values of ϵ and S can then be chosen for the beds of irregular particles so that the flow results lie on the same graph.

To chose a value for ϵ assume a value for the envelope density d of the granules (e.g. 0·7 g cm^{-3} for charcoal granules) and hence knowing the weight of the granules per unit volume of bed, ϵ can be determined. Once d is fixed the value of S is also fixed since d and S are interconnected through the following sequence of equations. (The terms used in the equations are defined after the complete sequence has been written down.)

$$k \text{ (Heywood's volume constant)} = 1/dNd_p^3. \qquad (2)$$

$$f \text{ (Heywood's surface coefficient)} = 1{\cdot}57 + Ck^{4/3}\{(n+1)/n^{1/3}\} \qquad (3)$$

$$S_0 \text{ (the specific surface determined optically)} = Nfd_p^2. \qquad (4)$$

Once S_0 is found S follows from the weight of unit volume of the granular bed.

In Eqns (2) to (4), N is the number of granules per gramme, n is the elongation index being the ratio L/B where L is the average length of the granules and B their average breadth (Heywood 1947), C is a constant put equal to 3·1, d_p 'the mean projected diameter' of a granule is defined by $d_p = 4A/\pi N$, where A is the specific projected area of the granules settled in their most stable position.

If an incorrect value is chosen for d in the first place then ϵ and S will be wrong and in the streamline region of flow the graph representing $\log \lambda$ as a function of $\log (Re)$ for the granules will be parallel to the graph established using lead shot. ϵ and S must now be altered to make the graphs coincide in the streamline region of flow. Consider the choice of d to be wrong by a factor x then k will be wrong by a factor $1/x$, f will be wrong by a factor y, where $y = \{1{\cdot}57 + Ck^{4/3}(n+1)/x^{4/3}n^{1/3}\}/\{1{\cdot}57 + Ck^{4/3}(n+1)/n^{1/3}\}$ and S_0 and S will both be wrong by a factor y. ϵ will be wrong by a factor z, where $z = (x - 1 + \epsilon)/x\epsilon$. Thus,

instead of $\log \lambda$ and $\log (Re)$ the wrong values $\log \lambda + \log z^3 - \log y$ and $\log (Re) - \log y$ have been calculated; in practice, $z^3 \gg y$, hence the displacement of the graph representing results for the granules is approximately $\log z^3$; thus z is found and, subsequently, the correct values of ϵ and S.

Experimental

The quantities in the equations (1) to (4) which are measurable are Δp, u, N, d_p and n. These were measured with apparatus very similar to that described in the previous paper (Jones 1956), using the same experimental technique; indeed, the only constructional alteration in the apparatus was in the flow apparatus in that alcohol manometers after the design of Kovačič (1953) were used for measuring pressure differences less than 2 cm w.g., instead of the aniline/sodium-carbonate manometer used previously. Two alcohol manometers were used; one measuring a pressure difference of 2 mm w.g. full scale, the other measuring 2 cm w.g. full scale.

The other quantities occurring in Eqns (1) to (4) are known from the known properties of air and the known dimensions of the charcoal beds, except of course, ϵ and S which are being determined. Having determined ϵ and S the quantities V (envelope volume per gramme) and S_0 could be calculated. V and S_0 should be constant for each charcoal independent of any variation in ϵ and S from bed to bed. The dimensions of the beds used in the experiments were determined by the dimensions of the canisters in which the beds were packed. Twelve canisters in all were used, comprising four canisters of each of the diameters 2, 3 and 4 cm. The four canisters of each diameter had lengths of 1, 2, 3 and 4 cm, respectively. More than one bed of any given dimension was made up during the course of the experiments. The beds were not used for the sole purpose of determining ϵ and S, but were used in uptake experiments; on the other hand, the flow characteristics of each was measured and the results can be used for the purpose of finding ϵ and S. For the largest granules flow experiments were also made with a bed 120 cm long and 2 cm in diameter.

In taking samples of charcoal for measurement of L and B care was taken to avoid choosing the larger particles because they are more evident in a number of granules spread on a white background; therefore, in sampling the charcoal, a spoonful was taken from the bottle and all the granules in the spoonful examined. Another spoonful was then taken and so on until at least 50 granules had been measured. The same procedure was followed in counting granules for the determination of N except that here the sample counted was about 0·5 gm. (In the case of the -52 $+60$ charcoal the sample was dispersed in a large known volume of glycerine and the number of granules in a known small part of this counted, after filtration.) To avoid dust adding to the value of A the granules were washed in alcohol before evaluating A. Five sieve fractions separated from one stock of charcoal were studied in the experiments. The charcoals are described in the Table.

Results obtained with flows of air from 0·03 l/min to 50 l/min through beds of lead shot are shown in Fig. 1(*a*). The straight portion of the curve corresponds to streamline flow through the beds and the point of departure of the curve from the straight line to the onset of turbulence; this occurs at a value of Reynolds number of about 3·0.

Several errors of observation can occur. At the smallest flow by far the largest are the errors in determining the pressure difference across the bed and in determining the flow itself; the other errors are negligible. In plotting the points on Fig. 1(*a*) at low flow (small values of Reynolds

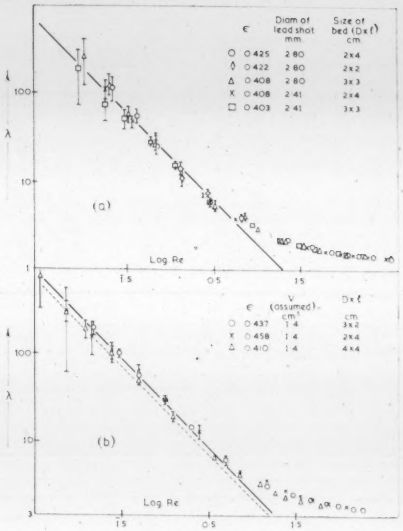


Fig. 1. Curves showing the relationship between the friction factor and Reynolds number for beds of (*a*) lead shot, (*b*) charcoal granules.

number), the most unfavourable case is represented, that is, where the two errors work in the same direction to produce a larger error. At high flows a number of observational errors of about the same magnitude occur; it is unreasonable to assume they can all act in the same direction so that the root mean square error is given for these.

The curve established in the streamline region of flow using lead shot beds is shown again in Fig. 1(*b*) by the broken line. The experimental points obtained using three different beds of charcoal from the B.S.S. range -8 $+10$ are shown on the figure. To calculate ϵ and S the superficial density was taken to be 0·715 g cm^{-3}; the best straight line through points at values of Reynolds number less than 3·0 is shown by the full line. This line represents streamline flow through the charcoal beds. The full line can be made to coincide with the broken line by taking the density of the charcoal to be 0·695 g cm^{-3}. Results for other grades of charcoals, together with shape factors relating to the charcoals are given in the Table. The symbols used in the Table are defined by equations (2) to (4).

The experimental errors in Fig. 1(*b*) are determined as for those in Fig. 1(*a*). It is important to note that care must be taken in measuring the diameter of the beds, since it ultimately occurs as the fourth power in λ and squared in (Re).

Discussion

It is known from other experiments that values determined for the surface area of granular particles depend on the experimental method used and, consequently, the surface area should be defined in relation to the use to which the values of the surface area are put. The work which has been described was carried out as part of a series of experiments to measure the uptake of benzene by charcoal beds as a function of the physical variables of the beds, the benzene

Table. *Description of the charcoals*

			Directly measurable					*Deduced from best fit in flow experiments*		
B.S.S. range	$\frac{D_m}{(cm)}$	N g^{-1}	A $(cm^2 g^{-1})$	dp (cm) ± 1 to 1%	L (cm) ± 1 to 2%	B (cm) ±2%	n ± 2 to 3%	$\frac{dp}{D_m}$	$V(cm^3 g^{-1})$ ± 1%	k ± 1 to 2%	$S_0(cm^2 g^{-1})$ ± 1%
− 8+10	0·187	318 ± 5	16·4 ± 0·4	0·256	0·303	0·210	1·44	1·37	1·44	0·269	57·0
−10+12	0·154	560 ± 6	19·7 ± 0·3	0·212	0·250	0·189	1·32	1·37	1·45	0·229	68·1
−14+16	0·110	1578 ± 9	27·3 ± 0·4	0·148	0·179	0·139	1·29	1·35	1·41	0·276	94·8
−18+22	0·078	4740 ± 25	41·8 ± 0·2	0·108	0·144	0·100	1·44	1·38	1·43	0·253	140·0
−52+60	0·027	$1·2 \pm 0·1 \times 10^5$	131 ± 2	0·0375	0·046	0·033	1·41	1·37	1·41	0·221	416

being carried in an air stream passing through the beds (Mills 1959), and it was thought that flow porosity and envelope surface area (as defined by flow) would be relevant variables to consider and therefore measure. But values of ϵ and S as defined by flow would be of value in many other problems where the flow of fluids around granular particles occurs, for example, not only in consolidated beds, but also in fluidized beds and under conditions where particles fall through fluids (Institute of Physics Conference 1951). The usefulness of determining ϵ and S (and hence V and S_0) by a flow method is therefore clear. The question arising is how far the method described here does measure the flow porosity and surface area of the granular beds. This question was in effect answered by the work of the previous paper (Jones 1956) when it was shown that Eqn (1) did apply to regular and irregular particles, the value of S for the irregular particles having been determined from the Eqns (2) to (4) and the envelope volume (and hence ϵ) from liquid displacement and a relationship between ϵ and D_m/D. In this work the envelope volume was not found from liquid displacement, but V and S_0 were found by trial and error by using the sequence of Eqns (1) to (4). This paper is, in fact, an account of the method of trial and error. As a check on this present work the volume of fluid displaced statically by the charcoal was found at the Chemical Defence Experimental Establishment, Porton, by immersing the unsieved charcoal in mercury (under a pressure of 1 atm) and also by immersing the charcoal in water, after treating the charcoal with silicone fluid. The purpose of the silicone treatment was to make the charcoal water-repellent so that water did not penetrate the interior of the granules. The specific volume V was found to be $1·38\ cm^3\ g^{-1}$ from immersion in mercury and also from immersion in water. On the other hand, the average value of V taken from the Table is $1·43\ cm^3\ g^{-1}$ which is $3\frac{1}{2}\%$ greater than the value found by liquid displacement. Since the value of V is the same by both mercury displacement and water displacement, the value 1·38 must be right since if there were much penetration into the interior it would be unlikely that it would be the same for the different liquids. The slightly greater volume found by air flow could be accounted for by supposing a stagnant layer of air of 10^{-4} cm thickness to cling to the particles during the flow experiments; considering the roughness of the particles such a stagnant layer would not be surprising. In any event it is felt the agreement between the results of the various methods of finding V, which have been described is sufficiently close to justify the use of air flow together with Heywood's shape factors to find porosities and surface areas for use in problems where fluid flow occurs.

Consider now the absolute position of the curve of Fig. 1(*a*) with respect to the coordinate axes. In the previous paper comparison was made between the results given there and those summarized by Rose and Mott (Institute of Physics Conference 1951), for regular particles, allowance being made for different definitions of λ and (Re), and it was seen that the results were in mutual agreement over those parts of the curves which were common to all investigations (values of $\log(Re) > 0·6$). The results given in Fig. 1(*a*) are similarly in agreement with previous results. Apart from the results of the previous paper and those given here no results are known for values of $\log(Re) < 0·6$; close examination of the results given here and those given in the previous paper will show they deviate from one another when $\log(Re) < 0·0$. Several precautions were taken in calibrating instruments for the determination of the absolute position of the line on Fig. 1(*a*) and its position is considered to be reliable; the linear part of the curve satisfies the equation $\lambda = 19·5/(Re)$. In streamline flow through straight tubes $\lambda = 16/(Re)$. The numerical constant 19·5 depends on definitions used in deriving expressions for λ and (Re) (cf. Rose (Institute of Physics Conference 1951), whose numerical constants are different from ours by a factor of 100), and also on the numbers substituted for quantities such as the tortuosity factor l_e/l; if in deriving Eqn (1) l_e/l were put equal to $\pi/2$ instead of $\sqrt{2}$ then $\lambda = 16/(Re)$ for beds as for tubes. The tortuosity factor was equated to $\sqrt{2}$ in Eqn (1) after Carman (1956, p. 45), but there is considerable justification for taking $\pi/2$ as the tortuosity factor in that it is the ratio between the semi-circumference and the diameter of a circle and thus represents the longer path taken round a sphere rather than through it (Carman 1956). To be able to represent flow through granular beds on the same curve as flow through tubes is a satisfying result and a pleasing application of the principles of similarity. To be able to do this directly Eqn (1) should be modified by putting $l_e/l = \pi/2$ instead of equal to $\sqrt{2}$; Eqn (1) then becomes

$$\lambda = 32\Delta p\epsilon^3 g/\pi^3 l\rho u^2\{S + (4/D)\}$$

and

$$R = \pi u\rho/\eta\{S + (4/D)\}. \qquad (5)$$

In conclusion, the absolute position of the curve of Fig. 1(*a*) does not affect the values found for ϵ and S for the charcoal beds. These values are chosen so that the curve representing the flow through the charcoal beds coincides with that representing the flow through the lead shot beds of known ϵ and S. Hence, constants in Eqn (1) and constants of calibration of flow apparatus do not alter the values found for ϵ and S of the charcoal beds.

The work was done in consultation with the Chemical Defence Experimental Establishment, Porton. We are particularly grateful to Dr. F. A. P. Maggs for his interest in the work.

References

CARMAN, P. C., 1956, *Flow of Gases through Porous Media* (London: Butterworths Scientific Publications).

HEYWOOD, H., 1947, Symposium on Particle Size Analysis, Supplement to *Trans Instn Chem. Eng.*, **25**, 18.

INSTITUTE OF PHYSICS, 1951, *Some Aspects of Fluid Flow* (London: Edward Arnold).

JONES, W. M., 1956, *Brit. J. Appl. Phys.*, **7**, 370.

KOVAČIČ, E., 1953, *J. Sci. Instrum.*, **30**, 304.

MILLS, D., 1959, Ph.D. Dissertation, University of Wales.

Measurement of thickness of thin transparent films using fluorescence

by B. P. BUNT, B.Sc., Research Department, The Metal Box Co. Limited, Acton, London, W.3

MS. received 30*th September* 1960, *in final form* 9*th December* 1960

Abstract

A molecular fluorescence method of comparing the thickness of transparent films (in the range 3×10^{-6} *to* 10^{-2} *cm) is described. An instrument for use in the middle of this range* (10^{-4} *cm) is also described and results show that its accuracy is to better than* $\pm 2\%$.

Introduction

MANY methods for measuring the thickness of films are available. When, however, a thin film is adhering to a relatively thick base considerable difficulties arise and it is frequently necessary to separate the film destructively from its substrate to carry out the measurement. This paper describes a non-destructive method of comparing the thickness of transparent or partially transparent films in the range 3×10^{-6} to 10^{-2} cm.

Principle

Molecular fluorescence has long been used as a means of chemical analysis and more recently it has been used for inspection purposes, e.g. crack detection. Sawyer and Heath* have described a quantitative method for air pollution studies in which a fluorescent dye is mixed with the pollutants and the distribution of the fluorescent tracer used to infer the distribution of the pollutant. This principle can be extended to the measurement of the thickness of films which are either fluorescent or which can be dyed with a fluorescent dye. Let us consider the system of Fig. 1, where we have a fluorescent film of thickness t on a non-fluorescent opaque substrate illuminated at an angle θ by light of wavelength λ_1 and intensity I. The fluorescent light of wavelength λ_2 and intensity V at an angle ϕ to the film is measured by the detector D. It is shown in the appendix that

$$V = \frac{KAIH}{4\pi Q}\left[\left\{1 + r\rho + \frac{Q}{T}(r-\rho)\right\}e^{-Qt} - r\rho e^{-2Qt} + \frac{Q}{T}re^{-(T+Q)t} + \frac{Q}{T}\rho e^{(T-Q)t} - 1\right]$$

where

$$Q = an_1(n_1^2 - \sin^2\theta)^{-\frac{1}{2}} + bn_2(n_2^2 - \sin^2\phi)^{-\frac{1}{2}}$$
$$T = an_1(n_1^2 - \sin^2\theta)^{-\frac{1}{2}} - bn_2(n_2^2 - \sin^2\phi)^{-\frac{1}{2}}$$

and K is the efficiency of the fluorescent conversion, A is the area of film viewed, H the solid angle of the cone of light entering the detector, r and ρ are the reflectivities of the substrate at λ_1 and λ_2, a and b are the absorption coefficients and n_1 and n_2 the refractive indices of the dyed film at λ_1 and λ_2.

This relation holds both for materials which are self-fluorescent and for those in which a fluorescent dye is uniformly dispersed. In general, the relation between thickness and measured fluorescent intensity is not linear, but where Qt and $Tt \ll 1$ the relation approximates to the linear. In the case of films which are non-fluorescent and sensibly transparent in themselves and to which a dye is added, a, b and K are functions of dye concentration. It is therefore possible to increase the range of thickness over which the approximation holds by decreasing the concentration of dye.

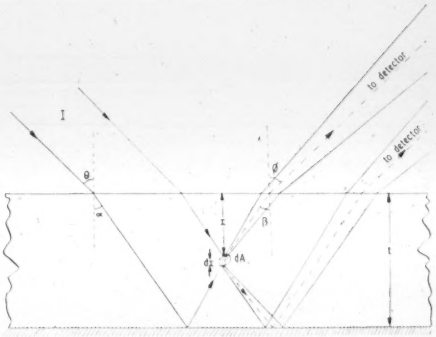


Fig. 1. Measurement of film thickness.

In this laboratory we have utilized this principle to measure the thickness of a variety of films on several different substrates. Most of our work has been with organic coatings and adhesives, of the order of 10^{-4} cm thick, on tin or aluminium substrates. For these films a fairly low dye concentration (0·03%) was shown to give adequate sensitivity with the instrument described below, and at this concentration the relation between thickness and fluorescent intensity was linear for all film thicknesses up to at least 10^{-3} cm. In some cases considerably thicker films have been successfully measured by this method. Polythene has been measured in thicknesses up to 10^{-2} cm and dioctyl sebacate films ranging in thickness from 7×10^{-3} cm down to an average of about 10^{-6} cm have been successfully measured. With the thinnest of these films it was, as expected, possible to preserve linearity with a considerably increased dye concentration. This ability to increase dye concentration in thin films greatly increases the range of the instrument to be described. If the condition that Qt and $Tt \ll 1$ is not

* Ministry of Supply, Chemical Warfare Defence Establishment, Porton Technical Paper No. 180, May 1950.

fulfilled the relation between I and t becomes more non-linear with increasing thickness and ultimately I becomes virtually insensitive to changes in thickness.

Apart from the need to exclude ambient light the method is non-contacting and may, therefore, be extended to the continuous measurement of the thickness of films on substrates which are in the form of moving webs. An instrument for this purpose is at present under development in this laboratory.

Apparatus

The fluorimeter consists of two units, the viewing head containing the light sources and the detector, and another unit which houses the photometer circuit and power supplies.

The viewing head is shown in Fig. 2. The head is divided into two compartments by the shelf A. The upper part

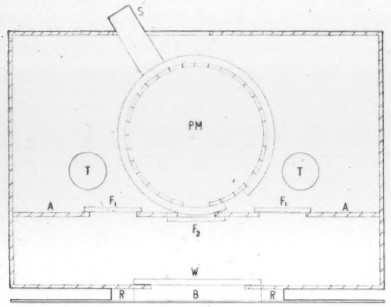


Fig. 2. The viewing head.

A, shelf supporting filters; T, ultra-violet emitting tubes; PM, side-window photomultiplier; S, shutter; B, sample; F_1, ultra-violet transmitting filters; F_2, ultra-violet stopping filter; W, quartz window; R, sponge-rubber framework.

contains two 4-w 6-in. fluorescent tubes T which have peak emission at a wavelength of about 360 mμ. The two Wratten 18A filters (Kodak Ltd.) F_1 allow only the ultra-violet light emitted by these tubes to pass into the lower compartment illuminating the sample at about 50°. A 27 M1 photomultiplier (Siemens Edison Swan Ltd.) PM is enclosed in a light-tight housing, the window of which opens into the lower compartment allowing the multiplier to view the sample normally. A light-tight shutter S is provided in front of this window, enabling the head to be moved without high light intensity damaging the photomultiplier. On the underside of the shelf there is an unmounted Wratten 2B gelatine filter F_2 in front of the photomultiplier window. Below the multiplier there is a quartz window W through which the sample is viewed. Outside the bottom of the box the window is surrounded with a sponge-rubber framework which forms a light-tight seal with the sample.

Although the photomultiplier is protected by a 2B filter the layout of the head is such that no ultra-violet light is specularly reflected on to this filter as, in common with all filters tried, it fluoresces to some extent. The inside of the viewing head is painted with a matt black paint which has been selected to be as nearly non-fluorescent as possible.

The measuring circuit and power supplies are conventional, the circuit being shown in Fig. 3. The combination of the constant voltage transformer and the corona discharge tube stabilizes the supply voltage for the dynode chain to better

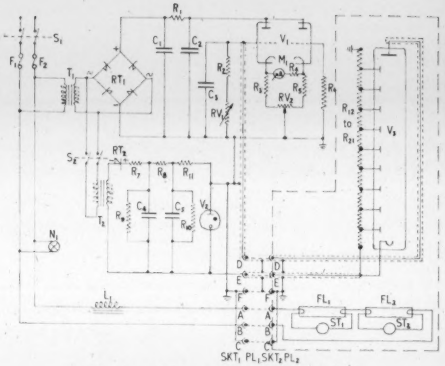


Fig. 3. Circuit diagram of fluorimeter.

R_1 = 22 kΩ ± 1%, ½ w H.S.
R_2 = 30 kΩ carbon resistor
R_3 = 6·2 kΩ carbon resistor
R_4 = 5 kΩ carbon resistor
R_5 = 4·3 kΩ carbon resistor
R_6, R_8 = 100 kΩ ± 10%, ½ w
R_7 = 47 kΩ carbon resistor
R_9, R_{10} = 10 MΩ carbon resistor
R_{11} = 470 kΩ carbon resistor
R_{12}–R_{21} = 150 kΩ ± 1%, ½ w H.S.

C_1, C_2 = 4 μF 450 v, d.c. electrolytic
C_3 = 0·1 μF 350 v, d.c. paper
C_{11}, C_5 = 0·25 μF 1500 v, d.c. paper

T_1 = 190–260 v/240 v, 15 w transformer

T_2 = 240 v/1000 v, 5 ma transformer

V_1 = twin triode
V_2 = stabilizer, 800 v
V_3 = photomultiplier

FL_1, FL_2 = 4 w, 6 in. u.v. tubes
ST_1, ST_2 = lamp starter
L_1 = lamp choke
M_1 = meter 0–100 ma
RT_1 = rectifier contact, full wave
RT_2 = rectifier, half wave
N_1 = neon indicator
F_1, F_2 = fuse 100 ma, 230/250 v a.c.
S_1, S_2 = on/off two pole 5 A

than 0·2%. The sum of the meter resistance and R_4 is such that the instrument has a linear response within the range of the meter. Changes in sensitivity are achieved by varying the anode load of the photomultiplier.

All thickness measurements made with this instrument are comparative, and a specially formulated uranium glass is used as standard. The instrument can be used for absolute thickness measurement, however, provided that a calibration experiment is carried out. When the film is dyed it is advisable to check the calibration for each batch of dyed film. Since the relation between fluorescence intensity and thickness is sensibly linear it is only necessary to check the calibration at a few points.

To take a reading, the viewing head is placed on a clean sample of the substrate and the zero set by adjusting RV_2. With the head placed on the standard, the upper end of the scale is then set by adjusting RV_1. The function of RV_2 is to allow for any asymmetry in the valve voltmeter circuit and to back-off both the photomultiplier dark current and the signal due to stray light. RV_1 varies the anode load of the multiplier and hence the sensitivity. This adjustment enables the reading for the standard to be set to a given value, thus compensating for changes in lamp brightness and other components. Usually a standard is chosen which fluoresces more brightly than the thickest sample to be measured. With some loss of accuracy samples which fluoresce more brightly than the standard can be measured by adjusting the sensitivity to give an intermediate reading on the standard.

The dye

Except in rare cases, where the coating has absorption and emission bands at appropriate wavelengths, it is necessary to add a fluorescent dye to the film. Most of the films examined

have been water white and for these a fluorescent brightening agent Tinopal PCRP (The Geigy Co.), which absorbs at around 360 mμ and emits at about 450 mμ was used. With the instrument described, only small amounts (0·03% has been found adequate) of this dye are required for film thicknesses of 5×10^{-4} cm. This material has the advantages that it is soluble in the solvents used in applying the films and is heat stable up to at least 300° C. Where the film is coloured, a dye should be chosen whose emission bands do not coincide with the absorption bands of the film.

The method can be extended to cases where the substrate is fluorescent. In this case it is necessary that the dye chosen should fluoresce at a wavelength well removed from that at which the substrate fluoresces. Suitable filters can then be used to absorb the unwanted fluorescence from the substrate. An alternative method where the substrate is uniformly fluorescent, is to use the decrease in apparent fluorescence of the substrate caused by absorption in the film. The brightness of the radiation reaching the detector is reduced if the film absorbs either the exciting or the emitted radiation. In certain types of film it is convenient to add an absorber of ultra-violet light while in other cases it is better to use a dye which absorbs the emitted light. The masking technique suffers from the minor defect that, if the film absorbs strongly, the relation between thickness and fluorescent intensity is non-linear while if the absorption of the film is feeble the thickness is found from the difference of two large quantities.

Results

Figure 4 illustrates the results obtainable with the instrument. Twenty-four small samples of tinplate were coated

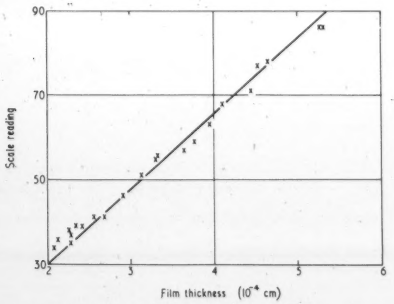


Fig. 4. Correlation of fluorescent reading with film thickness.

with a clear vinyl material containing 0·03% of dye. From a knowledge of the areas of the pieces of tinplate, their weights before and after coating and the density of the coating, the coating thickness on each can be calculated. Fluorescence intensity is expressed as a percentage of the brightness of the uranium glass standard. The correlation coefficient of film weight on fluorescence reading is 0·996. Instrumental stability was investigated by taking twenty readings of the fluorescence intensity of a block of uranium glass at three-minute intervals without altering either control. The standard deviation of these readings was 0·84% of the mean of the readings.

The instrument has been used to check the performance of machinery used to coat large sheets of tinplate. For this purpose it is necessary to determine film thickness at a considerable number of points on each of a number of consecutive sheets. The previously used method consisted of stamping out disks and weighing these before and after destructive removal of the coating. This procedure took several days to perform. Using the fluorimeter a more extensive survey is possible and the results can still be obtained in less than a day.

Acknowledgments

The author wishes to acknowledge his debt to Mr. M. J. Richard for discussions and advice on this work and to The Metal Box Co. Limited, for permission to publish this paper.

Appendix

The relation between fluorescent intensity and thickness

Let us consider an element V of volume $\delta A dx$ at a distance x below the surface of the film. Let us assume that after refraction at the surface of the film the illuminating beam makes an angle α to the vertical and the axis of the beam received by the detector makes an angle β to the vertical in the film, see Fig. 1.

Assuming the Bouguer absorption law, the intensity of incident light reaching A is given by

$$U = Ie^{-ax \sec \alpha} + Ire^{-(2t-x)\sec \alpha}.$$

The total fluorescent intensity emitted by the volume V is $F = KU\delta A dx$. If the solid angle, in the film, subtended by the detector is ϵ the total intensity of fluorescent light from V reaching the detector is

$$Z = \frac{\epsilon}{4\pi}\{Fe^{-bx \sec \beta} + F\rho e^{-b(2t-x)\sec \beta}\}.$$

The total intensity from the column of cross-sectional area δA reaching the detector is $Y = \int_0^t Z dx$. Hence

$$Y = \frac{I\delta AK\epsilon}{4\pi Q}\left[\left\{1 + r\rho + \frac{Q}{T}(r-\rho)\right\}e^{-Qt} - r\rho e^{-2Qt} - \frac{Q}{T}re^{-(r+Q)t} + \frac{Q}{T}\rho e^{(T-Q)t} - 1\right]$$

where $Q = a \sec\alpha + b \sec\beta$ and $T = a \sec\alpha - b \sec\beta$. Since $\sin\alpha = n_1^{-1} \sin\theta$ and $\sin\beta = n_2^{-1} \sin\phi$, these become

$$Q = an_1(n_1^2 - \sin^2\theta)^{-\frac{1}{2}} + bn_2(n_2^2 - \sin^2\phi)^{-\frac{1}{2}}$$
$$T = an_1(n_1^2 - \sin^2\theta)^{-\frac{1}{2}} - bn_2(n_2^2 - \sin^2\phi)^{-\frac{1}{2}}.$$

Assuming that the geometry of the detecting system is such that all areas such as δA have, on average, an equal effect, we may write

$$V = \frac{IAKH}{4\pi Q}\left[\left\{1 + r\rho + \frac{Q}{T}(r-\rho)\right\}e^{-Qt} - r\rho e^{-2Qt} + \frac{Q}{T}re^{-(T+Q)t} + \frac{Q}{T}\rho e^{(T-Q)t} - 1\right].$$

Conformal transformations by analogue computer

by E. J. JOSS and D. S. ROSS, The Royal College of Science and Technology, Glasgow

M.S. received 5th October 1960

Abstract

This paper describes the use of an analogue computer to generate certain conformal transformations of the circle, these being of use in the analysis of stress patterns in the neighbourhood of holes of various shapes. The relationship between the form of transformation and the shape produced is also discussed.

Purpose

EXPERIMENTAL work, using photoelastic methods, has been carried out on stress concentration in the neighbourhood of holes of various shapes in a plate under tension (Ross 1958). The analytical determination of the stresses is possible using Muskhelishvili's (1953) method (Godfrey 1956) if the shape of the hole can be expressed as a conformal transformation of the circle. It was desired to obtain the transformations for the shapes of hole investigated experimentally, so that experimental and analytical results could be compared.

Theory

The complex number $z = r\exp(i\theta)$ gives a circle on the Argand diagram for constant r and varying θ. The form of transformation used was a power series in z, starting with z^{-1}.

For square and cruciform holes the transformation was:

$$z' = z^{-1} + az^3 + bz^7 + cz^{11} + \ldots \qquad (1)$$

where a, b, c, ... are coefficients, numerically less than 1, which diminish as the power of z increases. The form is based on:

$$z' = \{z^{-1} + z^2\}^{\frac{1}{2}} \qquad (2)$$

which gives $a = \frac{1}{2}$, $b = -\frac{1}{8}$, $c = +\frac{1}{16}$.

If $z = r\exp(i\theta) = \cos\theta + i\sin\theta$ for $r = 1$

$$z' = \cos\theta + a\cos 3\theta + b\cos 7\theta + c\cos 11\theta + \ldots + i(-\sin\theta + a\sin 3\theta + b\sin 7\theta + c\sin 11\theta + \ldots) \qquad (3)$$

Computer set-up

Four Government-surplus sine–cosine potentiometers of the rectangular-card type were geared together with an arrangement of change wheels, so that different speed ratios were available (Fig. 1).

For the above transformation the gearing was arranged to give potentiometer speeds of ω, 3ω, 7ω and 11ω rad/s when the assembly was driven by an electric motor. The potentiometers were fed from a common $\pm$ 100 v supply and the phasing adjusted so that the voltages $100\cos\omega t$ $-$ $100\sin\omega t$, $100\cos 3\omega t$, etc., were produced (Fig. 2). Each voltage was then passed through a sign-reversing amplifier

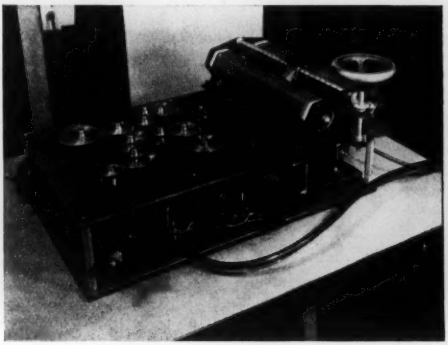


Fig. 1. Arrangement of change wheels allowing different speed ratios for potentiometers.

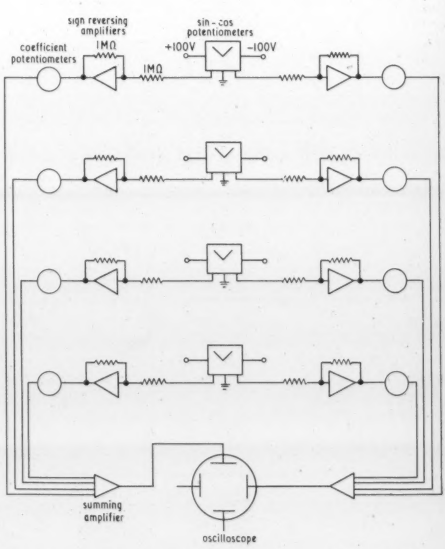


Fig. 2. Circuit diagram used for arrangement shown in Fig. 1.

to an adjustable ten-turn helical potentiometer, the reversing amplifier being used to reduce loading errors in the sin–cos potentiometers. The helical potentiometers could be set to give the required coefficients a, b, c, . . . with an accuracy to 0·1% of the full-scale travel of the potentiometer, using a bridge circuit to allow for loading effects. The resulting cosine and sine voltages were added in separate summing amplifiers, the sums being the x and y components of z'. These were fed to the x- and y-plates of an oscilloscope and, as the potentiometers rotated, the transformation was traced out on the oscilloscope screen.

Figs 3(*a*)–(*k*) show some of the results obtained.

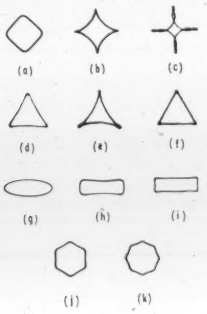


Fig. 3. Various results obtained on oscilloscope screen as the potentiometers are rotated.

(*a*) $z^{-1} + \frac{1}{6}z^3$. (*b*) $z^{-1} + \frac{1}{3}z^3$. (*c*) $z^{-1} + \frac{1}{2}z^3 - \frac{1}{8}z^7 + \frac{1}{16}z^{11}$.
(*d*) $z^{-1} + \frac{1}{2}z^2$. (*e*) $z^{-1} + \frac{1}{2}z^2$. (*f*) $z^{-1} + \frac{1}{4}z^2 + (1/45)z^5 + (1/162)z^8$.
(*g*) $z^{-1} + \frac{1}{2}z$. (*h*) $z^{-1} + \frac{1}{2}z - \frac{1}{8}z^3$. (*i*) $z^{-1} + \frac{1}{2}z - \frac{1}{8}z^3 - 0{\cdot}038z^5$.
(*j*) $z^{-1} + \frac{1}{16}z^5$. (*k*) $z^{-1} + \frac{1}{28}z^7$.

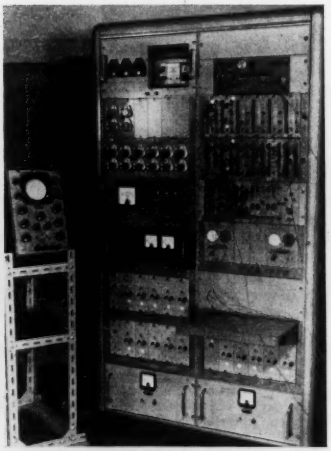


Fig. 4. A general view of the computer.

Choice of form of transformation

The first two terms in the series, $z^{-1} + az^3$, give the transformation its basically square shape, while the higher terms, having smaller coefficients have minor effects on the shape, being used mainly to adjust corner radii. The term z^{-1} can be regarded as a unit vector tracing a unit circle clockwise at ω rad/s, while the az^3 term gives a vector of length a rotating anticlockwise at 3ω rad/s. When $a = \frac{1}{3}$, the transformation is the hypocycloid generated by the circle radius a rolling inside a circle radius $1 + a$. This curve has four cusps and is shown in Fig. 3(*b*).

For small values of a the transformation is a 'square' with convex sides and large corner radii, and as a increases to $\frac{1}{3}$ the sides become concave and the corner radius decreases to zero. For $a > \frac{1}{3}$ loops appear at the corners of the figure.

Suitable choice of a, b, and c will result in the cruciform figure (Fig. 3*c*) with various radii at the ends of the arms.

The transformations $z^{-1} + az^2$, $z^{-1} + az^5$ will produce triangular and hexagonal forms respectively, the number of sides in the figure being one greater than the higher index of z.

Rectangular shape

This transformation includes a term in z which combines with the z^{-1} term to give an ellipse. This is then developed into a four-sided figure by adding a term in z^3.

$$\text{i.e.} \quad z' = z^{-1} + az + bz^3 + cz^5 + \ldots$$
$$= (1 + a)\cos\omega t + b\cos 3\omega t + \ldots$$
$$+ i\{-(1 - a)\sin\omega t + b\sin 3\omega t + \ldots\} \quad (4)$$

The first terms of the two series give an ellipse, axes $1 + a$ and $1 - a$.

Conclusion

The equipment described made exploration of the various transformations simple and rapid. The subsequent stress analysis based on the transformations obtained gave very satisfactory agreement with the experimental results.

The accuracy of the work is limited by errors in the sine–cosine potentiometers, the oscilloscope, and in measurement of the photographs. If greater accuracy had been desired the two latter sources of error could have been eliminated by taking static readings of the components of z', using the measuring bridge of the computer.

Acknowledgments

Thanks are due to Prof. A. S. T. Thomson, Head of the Department of Mechanical, Civil and Chemical Engineering, Royal College of Science and Technology, for the facilities provided in the Department laboratories, and to Mr. D. S. Ross, who provided the problem and arranged for the construction of the gearbox. The services of the technicians who constructed the equipment are also gratefully acknowledged.

References

GODFREY, D. E. R., 1956, *Report on Savin's 'Stress Concentration around holes', Part 1—Isotropic flat plates.* Ministry of Supply T.I.L. No. 18460.

MUSKHELISHVILI, N. I., 1953, *'Some basic problems of the mathematical theory of elasticity'*, translated from the Russian by J. R. M. Radok, (Gronigen: Noordhoff).

ROSS, D. S., 1958, *Photoelastic investigation of the stress concentration effects of discontinuities in tension plates.* (Unpublished paper given at the Conference of the Stress Analysis Group of The Institute of Physics, Glasgow, March 1958.)

Method for the consistent transformation of physical equations into non-dimensionally parametric form

by J. C. AMSON, B.Sc., British Welding Research Association, Abington, Cambridge

MS. received 26th September 1960

Abstract

Necessary and sufficient conditions for the existence of a simplifying transformation of a set of physical equations are given, whereby the physical variables involved become a less numerous set of dimensionless parameters. The proof provides a method for performing this transformation and for deciding whether it is unique, and this method is described. Whilst such a transformation is always possible for a single physical equation whose form is algebraic, its existence and description is not always apparent when more than one such algebraic equation is affected and the transformation is required to apply consistently to each, or when the equations are not algebraic. The method described is applicable not merely to sets of general equations, but to any collection of power-products of physical variables whose consistent transformation is desired.

1. Introduction

PHYSICAL equations often involve a large number of physical quantities as variables, many of which remain effectively constant during the description of a particular physical situation; the gravity acceleration variable, for example, is effectively a constant in most dynamical situations located at the earth's surface, yet appears prolifically in appropriate dynamical equations. The presence of such 'numerically static' variables, whilst physically necessary in any complete equation, detracts not only from the ease with which physical equations involving them can be manipulated, but also from the ease with which the 'working' variables are recognized as such so far as that particular physical situation is concerned. Methods of dimensional analysis (Bridgman 1931, Langhaar 1951) are known which (*inter alia*) condense automatically the number of variables by assimilating the numerically static variables into clusters about the working variables. These clusters are dimensionless and are, in fact, parameters which describe lucidly the behaviour of the physical situation concerned in terms of the non-transformed working variables. Moreover, these parameters are linearly independent or else linearly based on an independent set of dimensionless parameters, the so-called complete set (Birkhoff 1950, Langhaar 1951), and any non-dimensionally parametric description is always in terms of such a linearly based set. Now whenever the physical equations have been derived by dimensional methods, the choice of the dimensionless parameters as linear combinations of the complete set is clear; but the choice can be far from clear when we have to deal with a number of equations which have been derived by other than dimensional methods. Then the choice may be tedious or non-existent; that is, it might be impossible to choose a set of dimensionless parameters which will apply consistently to each of the given equations, and the lack of direct evidence of the impossibility cannot prevent much fruitless searching for the non-existent parameters.

But there are many advantages to be gained both physically and mathematically by the employment of dimensionless parameters in place of the actual physical variables, and it is often desirable to utilize fully these advantages without having to have obtained a set of equations via the dimensional method. The purpose of this paper is to give necessary and sufficient conditions for the existence of a consistent transformation of the physical equations, which will permit the description of the physical situation in terms of dimensionless parameters, and to describe a simple systematic method for performing the transformation whenever the conditions are satisfied in a particular case and the transformation has been shown to exist uniquely or otherwise.

2. The physical equations

We suppose initially that the physical equations we are to transform are simultaneous, dimensionally homogeneous and complete (Focken 1953), but are otherwise not specified.

Suppose the terms of our equations form a set $\mathbf{S}$ of (N) non-zero power-products, each of which is typically of the form

$$k_w \prod_{r=1}^{l} x_r^{a_{rw}} \prod_{s=1}^{m} y_s^{b_{sw}} \qquad w = 1, \ldots N \tag{1}$$

where the k_w are non-zero scalar (dimensionless) coefficients, the x_r are (l) primary* physical variables, dependent on each other and on the y_s, the y_s are (m) secondary* independent physical variables, the a_{rw} and b_{sw} are finite scalar exponents and w is an orderly subscripting of the power-products in $\mathbf{S}$.

Suppose further, without loss of generality, that each member of $\mathbf{S}$ is already dimensionless. (If the member originally had been an argument of some transcendental function, it would be already dimensionless, and if it had come from an algebraic equation it could have been made dimensionless, if not already so, in virtue of the hypothesis of dimensional homogeneity, simply by having first divided each such equation by any one of its non-zero terms.)

Suppose $\mathbf{D}_1$ and $\mathbf{D}_2$ are the $(q \times l)$ and $(q \times m)$ dimension

* It is important to appreciate that this segregation of physical *variables* into primary and secondary sets is not to be confused with the distinction between fundamental and derived *quantities* (Focken 1953) used in dimensional analysis and sometimes called the primary and secondary *quantities* respectively. Here the segregation is for descriptive convenience only; the primary variables are the 'working' variables of the Introduction, and the secondary variables are the 'numerically static'.

matrices (Langhaar 1951) of the x_r and y_s variables respectively, with respect to a set of (q) dimensionally fundamental unit quantities; and $\mathbf{F} = ((f_{sr}))$ is the transpose of a matrix of exponents.

3. The transformation

Let $\mathbf{T}$ be a transformation such that

$$\mathbf{T} : x_r \rightarrow \xi_r = x_r \prod_{s=1}^{m} y_s^{f_{sr}}, \; r = 1, \ldots l \tag{2}$$

and (i) the transformed variables ξ_r are dimensionless parameters, i.e.

$$\mathbf{D}_1 + \mathbf{D}_2\mathbf{F} = \mathbf{O} \tag{3}$$

(ii) for each $w = 1, \ldots, N$ we have

$$k_w \prod_{r=1}^{l} x_r^{a_{rw}} \prod_{s=1}^{m} y_s^{b_{sw}} = k_w \prod_{r=1}^{l} \xi_r^{a_{rw}}. \tag{4}$$

This transformation preserves the simultaneity of the original equations, since by (4) each term is replaced by a fully equivalent one, the dimensional homogeneity, since by (3) and (4) each original dimensionless term is replaced by a dimensionless product of dimensionless parameters, and the completeness, since this property is unaffected by a transformation such as $\mathbf{T}$. It also reduces the original set of $l + m$ variables to a set of l dimensionless parameters, consistently throughout $\mathbf{S}$. Thus the transformation $\mathbf{T}$ is of the kind we are seeking; we shall now give the conditions for its existence.

4. Existence conditions

Let $\mathbf{A} = ((a_{rw}))$ be the $(l \times N)$ matrix of primary exponents, of rank R, and let $\mathbf{B} = ((b_{sw}))$ be the $(m \times N)$ matrix of secondary exponents. Let $Rk(\mathbf{M})$, $Rk\begin{bmatrix}\mathbf{M}\\ \mathbf{r}\end{bmatrix}$, $Rk(\mathbf{M}, \mathbf{c})$ be the rank of a matrix $\mathbf{M}$, of a matrix $\mathbf{M}$ augmented by a row vector $\mathbf{r}$, and of a matrix $\mathbf{M}$ augmented by a column vector $\mathbf{c}$, respectively. Let $\mathbf{e}_{s*}$ be the sth row vector of the unit matrix $\mathbf{I}_m$, and $\mathbf{e}_{*\lambda}$ be the λth column vector of the unit matrix $\mathbf{I}_{(l-R)}$. Let $\mathbf{D}_1$ be partitioned as $[\mathbf{D}_{11}, \mathbf{D}_{12}]$, where $\mathbf{D}_{11}$ and $\mathbf{D}_{12}$ are $(q \times R)$ and $\{q \times (l - R)\}$ submatrices respectively.

Then necessary and sufficient conditions for a transformation $\mathbf{T}$ to exist uniquely are either

$$R = Rk(\mathbf{A}) = Rk\begin{bmatrix}\mathbf{A}\\ \mathbf{e}_{s*}\mathbf{B}\end{bmatrix}, \; s = 1, \ldots m, \; R = l;$$

or $$R = Rk(\mathbf{A}) = Rk\begin{bmatrix}\mathbf{A}\\ \mathbf{e}_{s*}\mathbf{B}\end{bmatrix}, \; s = 1, \ldots m, \; R < l;$$

and $$Rk(\mathbf{D}_2) = Rk\{\mathbf{D}_2, (-\mathbf{D}_{12}\mathbf{e}_{*\lambda})\} = m \quad \lambda = 1, \ldots l - R.$$

The necessary and sufficient conditions for an infinity of transformations $\mathbf{T}$ to exist, are

$$R = Rk(\mathbf{A}) = Rk\begin{bmatrix}\mathbf{A}\\ \mathbf{e}_{s*}\mathbf{B}\end{bmatrix}, \; s = 1, \ldots m, \; R < l,$$

and $$Rk(\mathbf{D}_2) = Rk\{\mathbf{D}_2, (-\mathbf{D}_{12}\mathbf{e}_{*\lambda})\} < m \quad \lambda = 1, \ldots l - R.$$

The necessary and sufficient condition for no transformation $\mathbf{T}$ to exist, is

$$R = Rk(\mathbf{A}) < Rk\begin{bmatrix}\mathbf{A}\\ \mathbf{e}_{s*}\mathbf{B}\end{bmatrix} \text{ for at least one value of } s, \quad s = 1, \ldots m.$$

5. Proof of existence conditions

We shall prove sufficiency only, whence, since all steps are reversible, the truth of the converse is immediate.

Whenever $Rk(\mathbf{A}) = Rk\begin{bmatrix}\mathbf{A}\\ \mathbf{e}_{s*}\mathbf{B}\end{bmatrix}$, $s = 1, \ldots m$

the equations $\mathbf{z}_s\mathbf{A} = \mathbf{e}_{s*}\mathbf{B}$

are soluble for $\mathbf{z}_s$, $s = 1, \ldots m$,

where $\mathbf{z}_s$ is a $(1 \times l)$ row vector.

Putting $\mathbf{z}_s = \mathbf{e}_{s*}\boldsymbol{\zeta}$, $s = 1, \ldots m$

we have $\mathbf{e}_{s*}\boldsymbol{\zeta}\mathbf{A} = \mathbf{e}_{s*}\mathbf{B}$, $s = 1, \ldots m$

from which we construct row-wise the equation

$$\boldsymbol{\zeta}\mathbf{A} = \mathbf{B} \tag{5}$$

which obviously retains solubility for $\boldsymbol{\zeta}$.

Hence the equation $$\mathbf{F}\mathbf{A} = \mathbf{B} \tag{6}$$

is soluble for $\mathbf{F}$, and we may re-write it in subscript notation,

as $$b_{sw} = \sum_{s=1}^{l} f_{sr}a_{rw}, \; w = 1, \ldots N.$$

Hence $$\prod_{s=1}^{m} y_s^{b_{sw}} = \prod_{s=1}^{m} y_s^{\left(\sum_{r=1}^{l} f_{sr}a_{rw}\right)}$$

and conversely, since the y_s are independent and non-zero by hypothesis. Thus

$$\prod_{s=1}^{m} y_s^{b_{sw}} = \prod_{s=1}^{m}\prod_{r=1}^{l} y_s^{f_{sr}a_{rw}} = \prod_{r=1}^{l}\prod_{s=1}^{m} y_s^{f_{sr}a_{rw}}.$$

By assumption the $a_{r\,w}$ are finite and $k_w \neq 0$, $w = 1, \ldots N$ and $x_r \neq 0$, $r = 1, \ldots l$,

hence we have, non-trivially

$$k_w \prod_{r=1}^{l} x_r^{a_{rw}} \prod_{s=1}^{m} y_s^{b_{sw}} = k_w \prod_{r=1}^{l} x_r^{a_{rw}} \prod_{r=1}^{l}\prod_{s=1}^{m} y_s^{f_{sr}a_{rw}}$$

i.e. $$k_w \prod_{r=1}^{l} x_r^{a_{rw}} \prod_{s=1}^{m} y_s^{b_{sw}} = k_w \prod_{r=1}^{l} \left\{x_r \prod_{s=1}^{m} y_s^{f_{sr}}\right\}^{a_{rw}} \quad w = 1, \ldots N.$$

Hence a transformation $\mathbf{T}'$ exists:

$$\mathbf{T}' : x_r \rightarrow \xi_r = x_r \prod_{s=1}^{m} y_s^{f_{sr}}, \; r = 1, \ldots l$$

which satisfies for each $w = 1, \ldots N$, the condition (4). We shall now show that a solution $\mathbf{F}$ of (6) is also a solution of (3) and hence that $\mathbf{T}' \equiv \mathbf{T}$.

Case (i): $R = Rk(\mathbf{A}) = l$.

By hypothesis there exists $\mathbf{A}'$, an $(l \times l)$ non-singular minor of $\mathbf{A}$.

Let $\mathbf{B}'$ be the $(m \times l)$ submatrix of $\mathbf{B}$, whose column subscripts are those of $\mathbf{A}'$. Let $\mathbf{a}_w$ and $\mathbf{b}_w$, $w = 1, \ldots N$, be the columns of $\mathbf{A}$ and $\mathbf{B}$ respectively.

Then since the terms (1) are dimensionless by assumption,

we have $$\mathbf{D}_1\mathbf{a}_w + \mathbf{D}_2\mathbf{b}_w = \mathbf{O}, \; w = 1, \ldots N$$

and we can construct column-wise the equation

$$\mathbf{D}_1\mathbf{A} + \mathbf{D}_2\mathbf{B} = \mathbf{O} \tag{7}$$

and in particular

$$\mathbf{D}_1\mathbf{A}' + \mathbf{D}_2\mathbf{B}' = \mathbf{O}, \quad |\mathbf{A}'| \neq 0. \tag{8}$$

Similarly by the construction of (5), we have from (6)

$$\mathbf{F}\mathbf{A}' - \mathbf{B}' = \mathbf{O} \quad |\mathbf{A}'| \neq 0. \tag{9}$$

Pre-multiplying (9) by $\mathbf{D}_2$

$$\mathbf{D}_2\mathbf{F}\mathbf{A}' - \mathbf{D}_2\mathbf{B}' = \mathbf{O}$$

and adding (8), we get

$$\mathbf{D}_2\mathbf{F}\mathbf{A}' + \mathbf{D}_1\mathbf{A}' = \mathbf{O}, \text{ i.e. } (\mathbf{D}_2\mathbf{F} + \mathbf{D}_1)\mathbf{A}' = \mathbf{O}$$

and since $|\mathbf{A}'| \neq 0$, we have finally

$$\mathbf{D}_2\mathbf{F} + \mathbf{D}_1 = \mathbf{O}.$$

Hence the solution $\mathbf{F}$ of (6) satisfies (3) as required, thus $\mathbf{T}' \equiv \mathbf{T}$. Further, the hypothesis $Rk(\mathbf{A}) = l$ is the condition for a unique solution $\mathbf{z}_s$ of $\mathbf{z}_s\mathbf{A} = \mathbf{e}_{s*}\mathbf{B}$, $s = 1, \ldots m$ and hence the condition for a unique solution $\mathbf{F}$ of (6). Since the transformation $\mathbf{T}$ is determined by $\mathbf{F}$, the uniqueness of $\mathbf{F}$ implies the uniqueness of $\mathbf{T}$.

Case (ii): $R = Rk(\mathbf{A}) < l$.

By hypothesis there exists an $(R \times R)$ non-singular minor of $\mathbf{A}$.

Suppose without loss of generality that

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{A}_{12} \\ \mathbf{A}_{21}\mathbf{A}_{22} \end{bmatrix}$$

such that $\mathbf{A}_{11}$ is of type $(R \times R)$, and $|\mathbf{A}_{11}| \neq 0$.

Since $\mathbf{F}$ is a solution of $\zeta\mathbf{A} = \mathbf{B}$ we can partition $\mathbf{F}$ and $\mathbf{B}$ so that

$$[\mathbf{F}_1\mathbf{F}_2]\begin{bmatrix} \mathbf{A}_{11}\mathbf{A}_{12} \\ \mathbf{A}_{21}\mathbf{A}_{22} \end{bmatrix} = [\mathbf{B}_1\mathbf{B}_2]. \tag{10}$$

Hence those equations involving the coefficients $\begin{bmatrix} \mathbf{A}_{12} \\ \mathbf{A}_{22} \end{bmatrix}$ are redundant.

Further, there are infinitely many solutions $\mathbf{F}$ of

$$[\mathbf{F}_1\mathbf{F}_2]\begin{bmatrix} \mathbf{A}_{11} \\ \mathbf{A}_{21} \end{bmatrix} = \mathbf{B}_1 \tag{11}$$

with $\mathbf{F}_2$ arbitrary.

Conformably partitioning $\mathbf{D}_1$ we have from (7)

$$[\mathbf{D}_{11}\mathbf{D}_{12}]\begin{bmatrix} \mathbf{A}_{11}\mathbf{A}_{12} \\ \mathbf{A}_{21}\mathbf{A}_{22} \end{bmatrix} + \mathbf{D}_2\mathbf{B} = \mathbf{O}. \tag{12}$$

Pre-multiplying (10) by $\mathbf{D}_2$, we have

$$\mathbf{D}_2[\mathbf{F}_1\mathbf{F}_2]\begin{bmatrix} \mathbf{A}_{11}\mathbf{A}_{12} \\ \mathbf{A}_{21}\mathbf{A}_{22} \end{bmatrix} - \mathbf{D}_2\mathbf{B} = \mathbf{O}. \tag{13}$$

Adding (12) and (13), we get

$$[\mathbf{D}_{11}\mathbf{D}_{12}]\begin{bmatrix} \mathbf{A}_{11}\mathbf{A}_{12} \\ \mathbf{A}_{21}\mathbf{A}_{22} \end{bmatrix} + \mathbf{D}_2[\mathbf{F}_1\mathbf{F}_2]\begin{bmatrix} \mathbf{A}_{11}\mathbf{A}_{12} \\ \mathbf{A}_{21}\mathbf{A}_{22} \end{bmatrix} = \mathbf{O}$$

from which we have in particular

$$\mathbf{D}_{11}\mathbf{A}_{11} + \mathbf{D}_{12}\mathbf{A}_{21} + \mathbf{D}_2\mathbf{F}_1\mathbf{A}_{11} + \mathbf{D}_2\mathbf{F}_2\mathbf{A}_{21} = \mathbf{O}$$

i.e. $$(\mathbf{D}_{11} + \mathbf{D}_2\mathbf{F}_1)\mathbf{A}_{11} + (\mathbf{D}_{12} + \mathbf{D}_2\mathbf{F}_2)\mathbf{A}_{21} = \mathbf{O}.$$

But $\mathbf{F}_2$ is arbitrary and by hypothesis

$$Rk(\mathbf{D}_2) = Rk\{\mathbf{D}_2, (-\mathbf{D}_{12}\mathbf{e}_{\lambda*})\}, \ \lambda = 1, \ldots l - R$$

so we can choose $\mathbf{F}_2$ such that

$$\mathbf{D}_{12} + \mathbf{D}_2\mathbf{F}_2 = \mathbf{O} \tag{14}$$

the choice being unique whenever $Rk(\mathbf{D}_2) = m$.

Then $$(\mathbf{D}_{11} + \mathbf{D}_2\mathbf{F}_1)\mathbf{A}_{11} = \mathbf{O}$$

but $$|\mathbf{A}_{11}| \neq 0,$$

hence $$\mathbf{D}_{11} + \mathbf{D}_2\mathbf{F}_1 = \mathbf{O}. \tag{15}$$

Eqns (14) and (15) are clearly equivalent to

$$\mathbf{D}_1 + \mathbf{D}_2\mathbf{F} = \mathbf{O}.$$

Hence the solution $\mathbf{F} = [\mathbf{F}_1\mathbf{F}_2]$ of (11) with $\mathbf{F}_2$ satisfying the condition (14), satisfies (3) as required, thus $\mathbf{T}' = \mathbf{T}$, $\mathbf{T}$ being unique or not, according as $Rk(\mathbf{D}_2)$ is equal or not to m.

Whenever $Rk(\mathbf{A}) < Rk\begin{bmatrix} \mathbf{A} \\ \mathbf{e}_{s*}\mathbf{B} \end{bmatrix}$ for at least one value of s, $s = 1, \ldots m$, then at least one equation $\mathbf{z}_s\mathbf{A} = \mathbf{e}_{s*}\mathbf{B}$, $s = 1, \ldots m$, is insoluble. Thus in this case, no equation $\zeta\mathbf{A} = \mathbf{B}$ can be constructed, hence an equation $\mathbf{F}\mathbf{A} = \mathbf{B}$ is insoluble for $\mathbf{F}$, and no transformation $\mathbf{T}$ exists. This concludes the proof.

6. Method for computing the transformation

The proof above provides a simple method for computing $\mathbf{T}$ whenever $\mathbf{T}$ exists.

If existence conditions (i) are satisfied, then $\mathbf{T}$ is found from (8), using

$$\mathbf{F} = \mathbf{B}\mathbf{A}^{-1}.$$

If existence conditions (ii) are satisfied, then $\mathbf{T}$ is found from (11), using

$$\mathbf{F}_1 = (\mathbf{B}_1 - \mathbf{F}_2\mathbf{A}_{21})\mathbf{A}_{11}^{-1}$$

and (14) $$\mathbf{D}_2\mathbf{F}_2 = -\mathbf{D}_{12}$$

the appropriate choices under (14) being dealt with as they arise.

Thus in either case the matrix $\mathbf{F}$ of the transformation is determined, and each column of $\mathbf{F}$ contains the exponents of the secondary variables which determine the transformation of the primary variable associated with that column. Then to effect the literal transformation $\mathbf{T}$ over $\mathbf{S}$, having recorded the above transformation of the primary variables, we may simply re-write every term of $\mathbf{S}$, retaining the scalar coefficients k_w unaltered, replacing each x_r by ξ_r to the same power and removing completely each y_s. The original set $\mathbf{S}$ is thus written consistently in non-dimensionally parametric form.

7. Example

The following example indicates the facility with which the transformation may be effected. A set of four equations:

$$\left.\begin{aligned} a^3\rho g &= k_1 b\tau - k_2 I^2, \ a > b \\ a^3 I^2 &= k_2 b^4\tau, \qquad a \leqslant b \\ fa^3\rho\epsilon &= k_3 Iv \\ fa^3 &= k_1 Wb^2 \end{aligned}\right\} \tag{16}$$

describes the consumption of a low-resistivity wire-rod anode in a vertical arc discharge, the anode becoming fused and

detached in a sequence of drops within the arc in a regular manner whilst being advanced towards the cathode at a regular rate. There are ten variables, five of which may be regarded as primary, i.e. $l = 5$, viz. W, wire-rod feed speed; f, frequency of drop detachment; I, arc current; a, radius of detached drop; b, radius of wire-rod anode; and five secondary, i.e. $m = 5$, viz. ρ, density of anode material; g, gravity acceleration; τ, surface tension of fused anode material; v, voltage drop across the arc system; ϵ, energy required to fuse unit mass of anode material.

To transform the set (16) into non-dimensionally parametric form, we first re-write them in dimensionless terms:

$$\left.\begin{array}{ll} 1 = k_1 a^{-3} b \rho^{-1} g^{-1} \tau - k_2 I^2 a^{-3} \rho^{-1} g^{-1}, & a \geqslant b \\ 1 = k_2 I^{-2} a^{-3} b^4 \tau, & a \leqslant b \\ 1 = k_3 f^{-1} I a^{-3} \rho^{-1} v \epsilon^{-1} & \\ 1 = k_1 W f^{-1} a^{-3} b^2 & \end{array}\right\} \quad (17)$$

There are five non-trivial terms, that is, $N = 5$. With the notation used above the exponent matrices are:

$$\mathbf{A} = \begin{array}{c} \\ W \\ f \\ I \\ a \\ b \end{array} \begin{array}{c} \begin{array}{ccccc} \text{Term number} & & & & \\ 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 2 & -2 & 1 & 0 \\ -3 & -3 & -3 & -3 & -3 \\ 1 & 0 & 4 & 0 & 2 \end{bmatrix} \end{array}$$

$$\mathbf{B} = \begin{array}{c} \\ \rho \\ g \\ \tau \\ v \\ \epsilon \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} -1 & -1 & 0 & -1 & 0 \\ -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{bmatrix} \end{array} \quad (18)$$

We shall form the dimension matrices $\mathbf{D}_1$, $\mathbf{D}_2$, with respect to the set of fundamental unit quantities: M (mass), L (length), T (time), as follows:

$$\mathbf{D}_1 = \begin{array}{c} \\ M \\ L \\ T \end{array} \begin{array}{c} \begin{array}{ccccc} W & f & I & a & b \end{array} \\ \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & \frac{1}{2} & 1 & 1 \\ -1 & -1 & -1 & 0 & 0 \end{bmatrix} \end{array}$$

$$\mathbf{D}_2 = \begin{array}{c} \\ M \\ L \\ T \end{array} \begin{array}{c} \begin{array}{ccccc} \rho & g & \tau & v & \epsilon \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & \frac{1}{2} & 0 \\ -3 & 1 & 0 & \frac{3}{2} & 2 \\ 0 & -2 & -2 & -2 & -2 \end{bmatrix} \end{array} \quad (19)$$

We may check Eqns (17), etc. from Eqn (7): $\mathbf{D}_1\mathbf{A} + \mathbf{D}_2\mathbf{B} = \mathbf{O}$. By inspection we note that

$$Rk(\mathbf{A}) = Rk \begin{bmatrix} \mathbf{A} \\ \mathbf{e}_{i*}\mathbf{B} \end{bmatrix} = 5,$$

and $R = l$; hence this is a sufficient condition for a unique transformation to exist, the transformation being determined by the matrix:

$$\mathbf{F} = \mathbf{B}\mathbf{A}^{-1} = \begin{array}{c} \\ \rho \\ g \\ \tau \\ v \\ \epsilon \end{array} \begin{array}{c} \begin{array}{ccccc} W & f & I & a & b \end{array} \\ \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{2} & \frac{1}{2} \\ -\frac{3}{4} & -\frac{5}{4} & \frac{1}{4} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} & -\frac{3}{4} & -\frac{1}{2} & -\frac{1}{2} \\ -1 & -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array} \quad (20)$$

from which we record the new form of the primary variables as dimensionless parameters:

$$\left.\begin{array}{rcl} W \to \Omega & = & W(\rho\tau g^{-3})^{\frac{1}{4}} \epsilon v^{-1} \\ f \to \Phi & = & f(\tau^3 \rho^{-1} g^{-5})^{\frac{1}{4}} \epsilon v^{-1} \\ I \to J & = & I(\rho g \tau^{-3})^{\frac{1}{4}} \\ a \to \alpha & = & a(\rho g \tau^{-1})^{\frac{1}{2}} \\ b \to \beta & = & b(\rho g \tau^{-1})^{\frac{1}{2}} \end{array}\right\} \quad (21)$$

That these parameters are in fact dimensionless is readily checked by inserting (19) and (20) into (3): $\mathbf{D}_1 + \mathbf{D}_2\mathbf{F} = \mathbf{O}$. Thus the unique non-dimensionally parametric form of the set of original equations (16) appears, under (21) as

$$\left.\begin{array}{ll} \alpha^3 = k_1\beta - k_2 J^2, & \alpha > \beta \\ J^2\alpha^3 = k_2\beta^4, & \alpha \leqslant \beta \\ \Phi\alpha^3 = k_3 J & \\ \Phi\alpha^3 = k_1\Omega\beta^2 & \end{array}\right\} \quad (22)$$

Acknowledgment

This paper was prepared at the British Welding Research Association, and my thanks are due to the Director, Dr. R. Weck, for his permission to publish.

References

BIRKHOFF, G., 1950, *Hydrodynamics: a Study of Logic, Fact and Similitude* (Princeton University).

BRIDGMAN, P. W., 1931, *Dimensional Analysis* (Yale University).

FOCKEN, C. M., 1953, *Dimensional Methods* (London: Arnold).

LANGHAAR, H. L., 1951, *Dimensional Analysis and the Theory of Models* (New York: John Wiley).

The determination of stress concentrations with an electrolytic tank model

by T. W. G. CALVERT, B.Sc.(Eng.), Grad.I.E.E.,* Imperial Chemical Industries Limited, Metals Division, Kynoch Works, Witton, Birmingham 6

MS. received 21*st February* 1960, *in revised form* 23*rd July* 1960

Abstract

In a loaded lamina of complicated shape stress concentrations arise which it is difficult to calculate. A model in the form of an electrolytic tank may be used to find the resultant stress at any point in a loaded lamina. The tank is made of an insulating material in the same shape as the lamina and filled with an electrolyte. It is fitted with electrodes to represent the areas over which the load is applied. The loads are represented in magnitude by the alternating currents flowing through the electrodes and in direction by the relative phase angles of the currents. The electric current at any point in the model tank can be sampled with a probe consisting of two fine wires a short distance apart. The resultant stress at a point in the loaded lamina is proportional to the electric current at the corresponding point in the model tank. This approach offers a valuable addition to means already available for finding stress concentrations in laminae of complicated shape.

Introduction

When a lamina of simple shape has a load applied to it, the form of the resulting stresses within it are usually obvious, and stress concentrations which result from changes in shape are calculable. When, however, the shape becomes more complicated, the stress distribution within the lamina is less obvious and calculations of stress concentrations become virtually impossible. In order to estimate stress concentrations, an empirical approach, such as the use of models, becomes necessary. The method commonly used employs the photoelastic properties of models made of a transparent plastic, but the analogy with the electric current distribution in an electrolytic tank also has much to recommend it.

Theory

This approach depends on an analogy between electrical and mechanical stress. It is well known that the general two-dimensional expression for the distribution of the sum of the principal stresses in an elastically strained lamina satisfies Laplace's equation, viz.

$$\frac{\partial^2 \Sigma}{\partial x^2} + \frac{\partial^2 \Sigma}{\partial y^2} = 0$$

where Σ = sum of principal stresses, i.e. the resultant stress at any point. Laplace's equation is also satisfied by the distribution of electric current in a lamina, hence

$$\frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} = 0.$$

In principle it should be possible to use an electric analogy based on the latter to represent the former, although the question of representing the boundary conditions may obviously present considerable problems. Such an analogy has in fact been used by Thum and Bautz (1934 see also Hetenyi 1950) to determine the stress distribution for a shaft of non-uniform cross section in torsion.

To see the application of the analogy to any two-dimensional system of forces acting on a lamina, consider first the simple case shown in Fig. 1. A homogeneous lamina of uniform thickness is represented by a similarly shaped tank containing

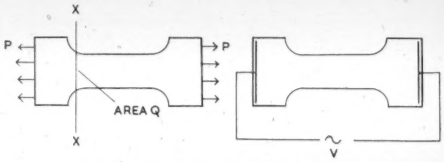


Fig. 1. A lamina with a uniformly distributed tensile load at each end is represented by a model tank with electrodes at each end.

a uniform depth of electrolyte. The uniformly distributed load on each end of the lamina is represented by an electrical potential applied to metal electrodes occupying each end of the tank. The resultant stress at any section of area is clearly P/a. The electrical analogy requires this to be represented by the current flowing in the section. It would be inconvenient to measure the current, but the potential gradient is proportional to it and can be found.

In the simple case considered, where there were only two points of load application, it followed as a condition of static loading that the two forces were equal and opposite. When there are three or more forces applied to a lamina in equilibrium, the algebraic sum of their resolved components in any given direction will be zero. If the current flow is to be analogous to the stress distribution, then the sum of the resolved components of the currents must also be zero. This can only be so if the current is alternating, and if the currents supplied to the electrodes differ in phase by angles corresponding to the angles between the lines of action of the forces they represent. Also the magnitudes of the currents must be in the ratio of the magnitudes of the forces they represent. Fig. 2 illustrates this in a simple case where forces are applied to the mid-points of the sides of a right-angled triangle. The potential gradient at any point in the electrolyte still represents the resultant mechanical stress at the corresponding point in the loaded lamina, but the potential gradient will now vary in phase throughout the tank. This

* Now at Canadair Limited, Montreal, Canada.

is not important since it is the magnitude of the potential gradient which gives the resultant stress.

A lamina of uniform thickness is easily represented by a tank with a uniform depth of electrolyte. A lamina of non-uniform thickness can be represented by a tank of non-uniform depth, if one face is plane and corresponds to the surface of the electrolyte. If one face is not plane, representation becomes difficult unless the lamina is symmetrical about some plane. This plane of symmetry can then be represented by the surface of the electrolyte.

As has been stated above, the current through each electrode should be proportional in magnitude and phase angle

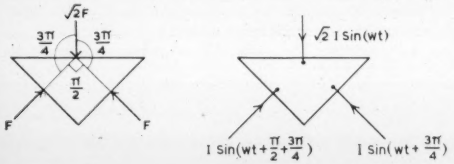


Fig. 2. The forces applied to the mid-points of the sides of a lamina in the shape of a right-angled triangle are represented by forces as shown.

to the magnitude and angle of application of the force which it represents. The area over which the force is applied is represented by the area of the electrode to which the corresponding current is fed.

Summarizing, it will be seen that boundary conditions can be satisfied such that any balanced two-dimensional system of concentrated or uniformly distributed forces acting on a lamina of uniform thickness can be represented. It is also possible to represent these forces applied to a lamina of non-uniform thickness if there is one plane surface or a plane of symmetry, since either of these can be represented by the surface of the electrolyte. In practice this means that there are very few problems which cannot be solved.

Electrolytic tank technique

General. Many problems, ranging from heat flow calculations to the design of electrode structures for electronic tubes, have been solved with the aid of an electrolytic tank analogy, and there exists in various published papers a considerable fund of experience on the technique of using the tank (see Boothroyd, Cherry and Maker 1948, Mickelsen 1949, Boothroyd 1951, Dadda 1951, Einstein 1951, Farr and Wilson 1951, Sanders and Yates 1952, Liebmann 1953, Diggle and Hartill 1954, Hartill, McQueen and Robson 1957). Since much of the experience is directly applicable to this use of a tank, only those aspects of tank technique where this method differs from others will be discussed in detail.

A problem encountered in all electrolytic tank work is the loss of accuracy due to the polarization potential drop at the surface of contact between the electrolyte and the electrodes. The errors are minimized by supplying the electrodes with alternating current of about 400 c/s. For optimum results the electrode material should have a fine grain structure and be coated with carbon or platinum black. The best electrolyte is distilled water, but tap water is satisfactory in most cases. The tank itself should be made or lined with a good insulating material and should be such that strips of plastic or wood can be temporarily cemented to it to build up the shape of the body in which it is desired to simulate the stresses.

Supply of current. The power capacity of the 400 c/s supply will vary with the size of tank and the total area of the electrodes. For a tank with its greatest dimension of the order of 4 ft, a supply of 25 watts has proved more than sufficient. Where it is only required to simulate two forces, the supply need have no adjustment. If more than two forces are required, but they are collinear, then it must be possible to vary and measure the current to each electrode. This can be done simply as shown in Fig. 3. The supply can be derived from a small alternator or an electronic oscillator.

In the general case where it is required to simulate more than two forces with any angle between them, it must be possible to independently measure and control the magnitude and phase of the current through any one electrode relative to the others. The exact form which the phase control takes depends on whether the supply is derived from an alternator or electronic oscillator. The method described makes use of a three-phase alternator.

The schematic diagram in Fig. 4 shows the general arrangement. The three-phase alternator has a tapped resistor between each phase, and other tapped resistors from these to the star point. R_A, R_B and R_C are used to give phase

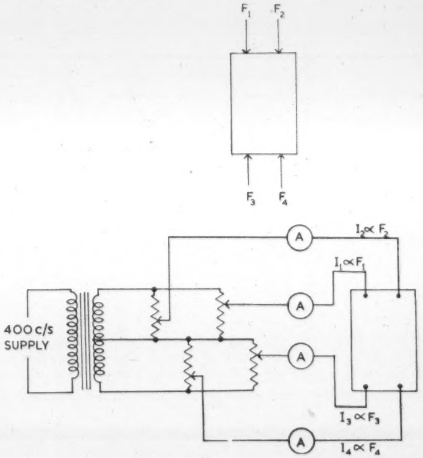


Fig. 3. For collinear forces the electrode currents are either in phase or 180 degrees out of phase, depending on whether the forces they represent reinforce or oppose each other.

variation and R_1, R_2, R_3 and R_4 give amplitude variation. Clearly one electrode supply will be used as a reference and fed direct from one of the phases (I_1). For each electrode, other than that used as a reference, a separate slider on R_A, R_B or R_C will be required. The position of the sliders on R_A, R_B and R_C can be calibrated to read off the relative phase angle directly. Then the amplitude of the current to each electrode can be set up by varying R_1, R_2, R_3 and R_4 and reading on the appropriate ammeter. Although this is not a particularly elegant or efficient method of obtaining the necessary variable phase supplies, it makes use of readily obtainable, robust units which are relatively immune from faults.

Measurement of potential. For the simpler cases, where there are only two forces or where there are more than two

collinear forces, the potential gradient at any point in the tank may be found by plotting the potential at a large number of points, and graphically determining the potential gradient

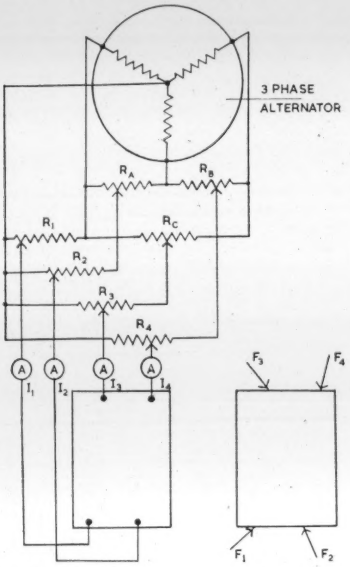


Fig. 4. The phase of the current to each electrode is varied by adjustment of a slider on R_A, R_B or R_C and the amplitude is varied by adjustment of R_1, R_2, R_3 or R_4.

(see Fig. 5). Even in simple cases this is tedious, and the coordinates of the points at which each measurement is made must be precisely recorded if the potential gradient is to be found accurately. This method is not easily applicable where the forces are not collinear since there is then a phase difference between the electrodes.

The potential gradient may be measured directly by using a probe which consists of two fine vertical wires fixed a

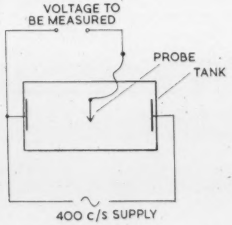


Fig. 5. A single wire probe may be used to plot the potential distribution.

small distance apart and just protruding into the electrolyte. By measuring the potential difference between these wires at any point, a measure of the potential gradient at that point and in the plane of the two wires is obtained. If the probe is made so that it can be rotated, then the maximum potential gradient at any point can be found. Recently an improved probe has been developed by Hartill, McQueen and Robson (1957) which eliminates the error due to the capillary action of electrolyte between the twin wires.

The potential across the probe may conveniently be measured with a potentiometer. The unknown potential is compared with a known potential and the difference indicated on a sensitive measuring device (e.g. a cathode-ray oscilloscope). The known potential is varied until a null point is reached, where the difference between the known and the unknown is a minimum. Fig. 6 shows the arrangement where there are two or more collinear forces, i.e. there is a single phase supply. The voltage applied to the electrodes

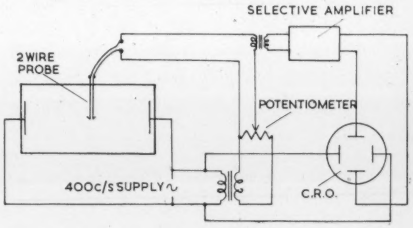


Fig. 6. A two-wire probe may be used to measure the potential gradient directly. The probe voltage is measured with a potentiometer.

is reduced by a transformer and connected across a potentiometer. The slider of the potentiometer moves over a scale and has the potential across the probe applied between it and one end of the potentiometer. The difference between the two potentials is developed across a small transformer and fed via an amplifier to a cathode-ray oscilloscope. The slider is adjusted so that there is a minimum deflection on the cathode-ray oscilloscope, i.e. until there is a horizontal line. The amplifier which feeds the error to the cathode-ray oscilloscope may be made sensitive only to the frequency used, and the effect of stray pick-up can thus be reduced.

If the forces represented are not collinear, the supply to the transformer feeding the potentiometer must be from a source which can provide a phase angle variation of 180 degrees. This can be arranged as shown in Fig. 7 where the three phase supply from the alternator is fed to a synchro-control transmitter (e.g. a Selsyn generator), where the rotor will develop a voltage, the phase of which is proportional to the angle of rotation of the rotor. The output from the rotor is fed to the potentiometer. To find the null point it is now necessary to successively adjust the phase-shift unit and the potentiometer so as to give the reference voltage the same phase and magnitude as the unknown voltage across the probe. When the phase-shift is out of adjustment the trace on the cathode-ray oscilloscope will open out into an ellipse, and when the potentiometer slider requires adjustment the trace will be diagonal instead of horizontal. For most applications the phase angle of the unknown probe voltage is not of interest, but it must be the same as the reference if the true magnitude of the probe voltage is to be measured. At each point of measurement the probe must be rotated until a maximum is found and in this position the probe wires will be normal to a line joining points of equal resultant stress.

If the model tank represents a lamina of non-uniform thickness and thus has a non-uniform depth, it may be

required to find the stress distribution on some plane other than that represented by the surface of the electrolyte. This is possible by inserting the probe to the required depth, but special attention must be given to probe design, since if the electrolyte is unduly disturbed inaccuracies will result.

Interpretation of results. It has been shown that a measure of the potential gradient at any point in a model may be obtained for almost any two-dimensional system of forces applied to the body it represents. To obtain the stress distribution, readings of potential gradient should be plotted on a diagram and equal points joined together by lines. If enough readings are taken a complete pattern of lines of equal resultant stress may be built up. However, this process which can be laborious, is seldom necessary, since with any specific problem it is generally known where the maximum stresses will occur and a solution can be quickly obtained with a few well chosen readings.

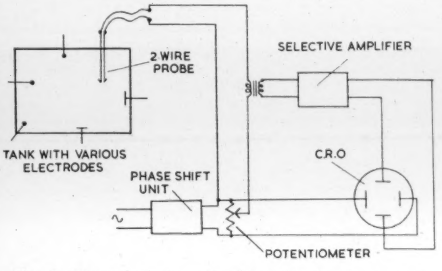


Fig. 7. The use of a two-wire probe when the forces represented are not collinear. The potentiometer must be fed from a variable phase supply.

Frequently the resultant stress in one part of the loaded lamina is easily calculable. When this is the case, the resultant stress in any other point in the lamina can be found by multiplying by the ratio of the potential gradients measured at corresponding points in the model tank.

If there is no point at which the resultant stress is calculable it is still possible to find the resultant stress at any point in the lamina, but the results are less accurate. It is first necessary to make an absolute measurement of the voltage across the probe and to know the relationship between the current through each electrode and the load it represents.

Assuming the model tank is full size, let an applied load be P kg and the corresponding current I A. Let the spacing between the probe wires be d cm and the resistivity of the electrolyte ρ ohm cm. Now suppose that at any point the unknown resultant stress is S kg/cm^2 and the measured probe voltage V. Then the current density at that point is $V/\rho d$ A/cm^2.

Since $$P \text{ kg} \equiv I \text{ A}$$

$$S = \frac{P}{I}\frac{V}{\rho d} \text{ kg/cm}^2.$$

Clearly it is not necessary for the model tank to be full size and simple scaling factors will allow the model to be made smaller or larger than the loaded lamina.

The accuracy of the results obtained by this method will depend on how accurately the probe voltage can be measured and how accurately the resistivity of the electrolyte and the distance between the probe wires is known. Because of this, the method is inferior to that demonstrated above which depends on a reference stress at some point in the loaded lamina being easily calculable.

Comparison with other methods. There are two other methods which are used to obtain the stress distribution in loaded laminae. The first involves a mathematical approach which by a process of successive approximations gives the value of the stress at any point. However, the process is tedious and must be repeated for any change in the load conditions. For these reasons, and since the results obtained are not always reliable, this method is little used except in simple cases. The second method, which is widely used, makes use of the photoelastic properties of a transparent plastic model (Heywood 1952), and it is worth considering in some detail the relative merits of this method and the electric analogy.

Each method depends on a model. The electric model is a tank filled with an electrolyte, and some work will be involved in building up the shape of the lamina to be simulated. The photoelastic model on the other hand, requires a careful selection of material, machining the material to shape, grinding the faces parallel, and annealing to remove any stresses. Generally each change in shape will require a fresh photoelastic model, while a tank can have its shape modified. Both methods require complicated auxiliary apparatus. For the electric analogy, a small alternator, and facilities for measuring the probe voltage are required. To make the photoelastic stress visible a polariscope is required, together with a loading frame to apply loads to the model and a photographic camera to record the results. An advantage of the electric analogy is that the loading conditions on a given model can be changed by the adjustment of control resistors, as opposed to extensive alterations to the loading frame to change the loads applied to a photoelastic model.

Both methods are most easily applied to simulate a two-dimensional system of forces acting on a plane lamina of uniform thickness. The electric analogy can be easily extended to cover cases where only one surface of the lamina is plane or where there is a plane of symmetry. The photoelastic method can be less easily extended to cover variation in thickness and three-dimensional loading by using a frozen stress technique. A major difference between the methods is that while the electric model cannot change shape when loaded, the photoelastic model will tend to deform in the same way as the lamina it represents. This can be an advantage, but inaccuracies may result unless the material of both the model and the lamina it represents have the same value of Poisson's ratio.

The above comparison shows that most stress distribution problems can be solved as easily and as cheaply with an electrolytic tank as with a photoelastic model. It is believed that in cases where a lamina of fixed shape is to be tested under widely different loading conditions the electric analogy is definitely preferable.

Conclusion. It has been shown that the electrolytic tank can be used to determine the stress distribution in a loaded lamina and to thus determine the values of the stress concentrations which may occur at points in the lamina. Where the stress in one part of the lamina is known, the stress in any other part may be determined accurately; where the stress in no part of the lamina is known or easily calculable, the stress in any part may still be determined, but to a lesser degree of accuracy since the result will depend on the accuracy to which the tank parameters are known. While the apparatus required to supply the current is undoubtedly complicated, it is simple to operate. For a given shape of model a new set of loading conditions may be set up very quickly. For a

new shape of model rebuilding is required but this need not be slow using flexible plastic materials and quick setting cements.

As a method of finding stress distribution this is compared with that which makes use of the photoelastic properties of transparent plastics. It is concluded that as a general tool this method is equally useful, and that for certain purposes it has definite advantages.

References

BOOTHROYD, A. R., 1951, *Proc. Instn Elect. Engrs*, **98,** 486, 65.

BOOTHROYD, A. R., CHERRY, E. C., and MAKER, R., 1948, *Proc. Instn Elect. Engrs*, **96,** 790, 163.

—— 1949, *Proc. Instn Elect. Engrs*, **97,** 126.

DADDA, L., 1951, *Energia Elett.*, **28,** 23.

DIGGLE, H., and HARTILL, E. R., 1954, *Proc. Instn Elect. Engrs*, **101,** 349.

EINSTEIN, P. A., 1951, *Brit. J. Appl. Phys.*, **2,** 49.

FARR, H. K., and WILSON, W. R., 1951, *Trans. Amer. Inst. Elec. Engrs*, **70,** 1301.

HARTILL, E. R., MCQUEEN, J. G., and ROBSON, P. N., 1957, *Proc. Instn Elect. Engrs*, **104,** 401.

HETENYI, M., 1950, *Handbook of Experimental Stress Analysis*, 1st Edn (London: Chapman and Hall), p. 748.

HEYWOOD, R. B., 1952, *Designing by Photoelasticity* (London: Chapman and Hall).

LIEBMANN, G., 1953, *Brit. J. Appl. Phys.*, **4,** 193.

MICKELSEN, J. K., 1949, *Gen. Elect. Rev.*, **52,** 19.

SANDERS, K. F., and YATES, J. G., 1952, *Proc. Instn Elect. Engrs*, **100,** 167.

THUM, A., and BAUTZ, W, 1934, *Verein Deutcher Ingenieure*, **78,** 17.

The coupling factor of piezoelectric ceramic disks

by D. S. CAMPBELL, B.Sc., and A. M. MACSWAN, Ph.D., The Plessey Co. Limited, Research Laboratories, Caswell, Towcester, Northants.

MS. received 28th September 1960

Abstract

Numerous formulae have been given for the calculation of the electromechanical coupling factor of piezoelectric disks in terms of the resonant and anti-resonant frequencies. These formulae, which are derived either on the basis of an analysis of the mechanical motions of a disk, or by considering the equivalent electrical circuit give widely divergent values of coupling factor.

It is shown that when an equivalent circuit which takes account of the overtones is used, the corresponding value of coupling factor obtained is in agreement with that of the mechanical analysis.

Curves, and a nomogram are given to enable coupling factor to be quickly calculated from the equations of the mechanical analysis without approximation.

1. Introduction

A PIEZOELECTRIC ceramic disk, silvered on both flat faces, will oscillate in several different modes if an alternating field is applied across the disk. The primary vibration will be in the thickness direction (i.e. parallel to the axis of the disk) and this will be due to the piezoelectric nature of the material. However, coupled to this vibration will be secondary vibrations, one of the more important of which is that in the radial direction.

It is possible for any of these modes to be in resonance and this paper is concerned with the behaviour of a disk when vibrating at or near a resonance in the radial mode.

When the impedance of a disk, in which the radius is very much greater than the thickness, is measured as a function of frequency it is found that the general trend is a decrease of impedance with frequency. In the appropriate frequency range, and for a disk 1 cm in diameter this will be between 100 kc/s and 1 Mc/s and there will be superimposed on the general decrease in impedance pairs of minima and maxima. These are known as resonance and anti-resonance respectively, the lowest in frequency corresponding to the fundamental radial mode, and those higher the first, second, and so on overtones. It is usually possible to measure up to three overtones on a normal ceramic disk. Fig. 1 shows

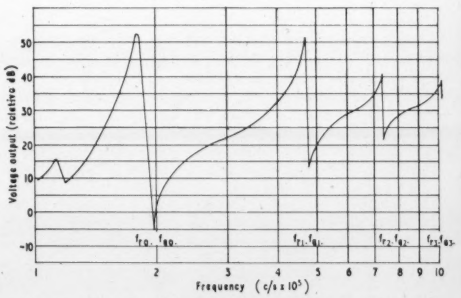


Fig. 1. Variation of voltage with frequency across a small resistor in series with a circular disk (1·3 cm diameter).

the variation of voltage with frequency across a small resistor (10 Ω) in series with a disk.

Since the material of the disk is piezoelectric, some of the input electrical energy is stored as mechanical energy. The coefficient of electromechanical coupling k can be defined at constant stress by the expression (see, for example, Katz 1959)

$$k^2 = \frac{\text{stored mechanical energy}}{\text{total stored energy}}.$$

With this definition, it is possible to derive formulae giving k^2 in terms of the resonant and anti-resonant frequencies at one particular resonance, on either an electrical or mechanical model. The resulting formulae are found not to be the same.

The purpose of this paper is to discuss this discrepancy and introduce modifications that reconcile the two approaches.

2. Mechanical analysis

A theoretical study of the behaviour of a piezoelectric disk polarized perpendicular to its flat surfaces has been given by Mason (1948). This analysis is confined to the radial vibrations of the disk: the following expression for the admittance of the disk is obtained.

$$Y = \frac{j\omega\pi a^2}{Kt}\left\{1 + \frac{k^2}{1-k^2}\,\frac{(1+\sigma)J_1(\omega a/v)}{(\omega a/v)J_0(\omega a/v) - (1-\sigma)J_1(\omega a/v)}\right\} \tag{1}$$

where $\omega = 2\pi \times$ frequency, $a =$ disk radius, $K =$ constant (function of permittivity), $t =$ disk thickness, $\sigma =$ Poisson's ratio, $v =$ velocity of sound in the disk material and J_0 and J_1 are Bessel functions of zero and first order.

Examination of the admittance expression shows it to be infinite (i.e. zero impedance or resonance) when

$$\frac{\omega a}{v}J_0\left(\frac{\omega a}{v}\right) = (1-\sigma)J_1\left(\frac{\omega a}{v}\right)$$

i.e.
$$\frac{\omega a}{v}J_0\left(\frac{\omega a}{v}\right)\Big/J_1\left(\frac{\omega a}{v}\right) = 1 - \sigma. \tag{2}$$

This defines a value of $\omega a/v$ which is a function of σ only.

The condition for anti-resonance is given by the vanishing of admittance

$$\frac{\omega a}{v}J_0\left(\frac{\omega a}{v}\right) = \left\{(1-\sigma) - (1+\sigma)\,\frac{k^2}{1-k^2}\right\}J_1\left(\frac{\omega a}{v}\right)$$

i.e.
$$\frac{\omega a}{v}J_0\left(\frac{\omega a}{v}\right)\Big/J_1\left(\frac{\omega a}{v}\right) = 1 - \sigma - (1+\sigma)\frac{k^2}{1-k^2} \tag{3}$$

from which is obtained a value of $\omega a/v$, this time corresponding to anti-resonance, which is a function not only of σ but also of k^2.

It is possible to use Eqns (2) and (3) to obtain the value of the coupling factor. If R_0 is the value of $\omega a/v$ corresponding to fundamental resonance, i.e. the first root of Eqn (2), then the Bessel functions $J_0(\omega a/v)$ and $J_1(\omega a/v)$ may be expressed by a Taylor expansion about R_0 and their values at anti-resonance determined. When this is done, it is found that

$$\frac{\Delta f}{f_r} = \frac{(1+\sigma)k^2/(1-k^2)}{R_0^2 - (1-\sigma^2) + (1+\sigma)k^2/(1-k^2)} \tag{4}$$

which Mason approximates to

$$\frac{\Delta f}{f_r} = \frac{(1+\sigma)k^2)(1-k^2)}{R_0^2 - (1-\sigma^2)}$$

i.e.
$$k^2 = \left(\frac{R_0^2 - (1-\sigma^2)}{1+\sigma}\,\frac{\Delta f}{f_r}\right)\Big/\left(1 + \frac{R_0^2 - (1-\sigma^2)}{1+\sigma}\,\frac{\Delta f}{f_r}\right). \tag{5}$$

From this equation k^2 may be calculated in terms of $\Delta f/f_r$, R_0 and σ. When $k^2 \ll 1$, Eqn (5) approximates to

$$k^2 = \frac{R_0^2 - (1-\sigma^2)}{1+\sigma}\,\frac{\Delta f}{f_r}. \tag{6}$$

Thus k^2 may be expressed in terms of f_r, f_a, R_0 and σ, and it is necessary to determine both σ and R_0.

From Eqn (2) it is possible to determine the value of the roots R_0, R_1, etc., as a function of σ only. Since $R_0 = \omega_{r0}a/v$ and $R_1 = \omega_{r1}a/v$ therefore

$$R_1/R_0 = f_{r1}/f_{r0}.$$

Thus a plot of σ against f_{r1}/f_{r0} can be used to determine σ. Fig. 2 shows the calculated curves for R_1/R_0 and R_2/R_0.

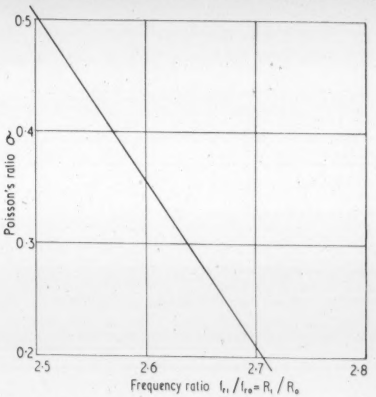


Fig. 2(*a*). Poisson's ratio, σ for varying values of the ratio of roots R_1/R_0.

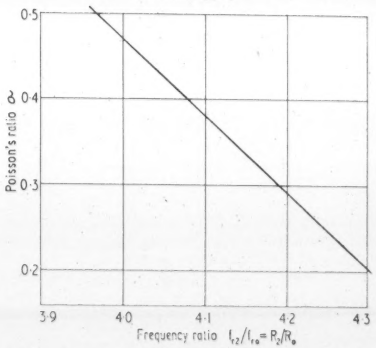


Fig. 2(*b*). Poisson's ratio, σ for varying values of the ratio of roots R_2/R_0.

Bogdanov and Timonin (1957) have derived a numerical expression for the first of these curves.

It is, however, possible to obtain a value of the coupling coefficient from the mechanical analysis without recourse to any of the above approximations. It will be noted that both Eqns (2) and (3) are of the same form, $xJ_0(x)/J_1(x) = A$, say. If this is plotted against x for a range of values in the region of resonance and anti-resonance (Fig. 3) the coupling coefficient may be determined. The value of Poisson's ratio is obtained as described previously and the corresponding value of R_0 at f_{r0} found. Then the value of the first root at anti-resonance R_0' may be obtained, since $R_0'/R_0 = f_{a0}/f_{r0}$ and from Eqn (3)

$$\frac{R_0'J_0(R_0')}{J_1(R_0')} = 1 - \sigma - (1+\sigma)\frac{k^2}{1-k^2}.$$

The value of the left-hand side of this equation, A, can be found from Fig. 3. Then since A and σ are known, the corresponding value of k^2 may be found from the nomogram (Fig. 4).

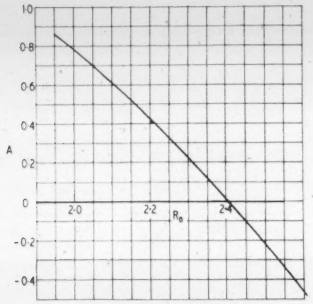


Fig. 3. Curve of A against R_0.

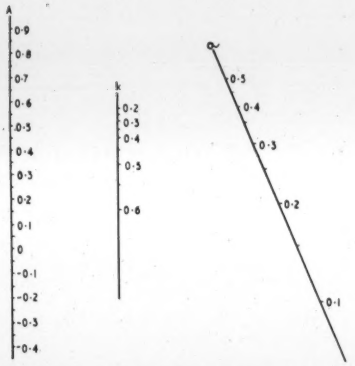


Fig. 4. Nomogram for determining coupling factor from known values of σ and A.

This procedure may also be carried out for the first and subsequent overtones and the appropriate graph of A against R_1 is given in Fig. 5.

3. Simple equivalent circuit

It is possible to simulate the behaviour of the disk by means of an equivalent electrical circuit. Although the present circuit is inadequate in certain respects which will be discussed in § 4, it is useful in the neighbourhood of a given resonance. The circuit is the same as that proposed for quartz by Van Dyke (1925) (Fig. 6).

L, C_1 and R arise from mechanical considerations, L being a function of mass, C_1 of the rigidity of the material and R being the resistance or internal friction. C_0 is the electrostatic capacity of the disk. The total impedance of the disk is then given by

$$Z = R + j\frac{\left[\dfrac{\omega^2LC_1 - 1}{\omega C_1}\left\{1 - \dfrac{C_0}{C_1}(\omega^2LC_1 - 1)\right\} - \omega C_0R^2\right]}{\left\{1 - \dfrac{C_0}{C_1}(\omega^2LC_1 - 1)\right\}^2 + \omega^2C_0^2R^2} \qquad (7)$$

In the case of $R = 0$, i.e. infinite mechanical Q_m, the relationship between resonance ($Z = 0$) and anti-resonance ($Z = \infty$) is given by

$$\omega_a^2 = (1 + C_1/C_0)\omega_r^2$$

i.e. $$\frac{\omega_a^2 - \omega_r^2}{\omega_r^2} = \frac{C_1}{C_0} \quad \text{or} \quad \frac{2\Delta\omega}{\omega_r} \simeq \frac{C_1}{C_0}.$$

The electromechanical coupling coefficient is defined as the ratio of stored mechanical energy to total stored energy and the energy is taken as proportional to the appropriate capacitor value. Thus,

$$k^2 = \frac{C_1}{C_1 + C_0} \simeq \frac{2\Delta f/f_r}{1 + 2\Delta f/f_r} \qquad (8)$$

so that again, a measurement of f_r and f_a will give the coupling coefficient.

When $\Delta f/f_r \ll 1$ (i.e. $k^2 \ll 1$) the approximate formula

$$k^2 = 2\Delta f/f_r \qquad (9)$$

will hold. This form is widely used in the measurement of coupling factor.

This may be compared with the corresponding value from Mason's equation (6) and it will be noted that the factor $\{R_0^2 - (1 - \sigma^2)\}/(1 + \sigma)$ is replaced by 2. For $\sigma = 0{\cdot}27$, $R_0 = 2{\cdot}03$ and this factor has the value $2{\cdot}51$. Thus, there is a considerable discrepancy between the results of the mechanical and electrical analyses as so far developed.

The effect of dielectric loss can be represented by a parallel resistance across C_0. Although this will alter the expression for Z, for all practical materials the effect on the resonant and anti-resonant frequencies is found to be negligible.

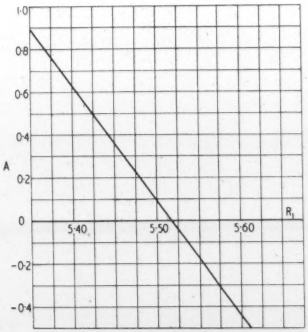


Fig. 5. Curve of A against R_1.

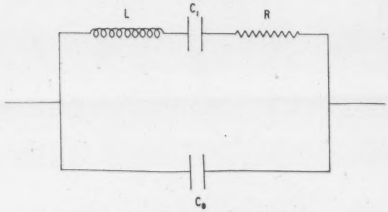


Fig. 6. Van Dyke equivalent circuit.

Dependence of circuit parameters on disk dimensions

In the mechanical analysis it is shown that f_r and f_a are functions of the radius only while the electrostatic capacity depends on both radius and thickness. The effect of thickness on the mechanical elements in the equivalent circuit may be deduced from the following observed phenomena: (i) the coupling factor k is a property of the material, hence C_1/C_0 is a constant; (ii) the resonance behaviour of a given mode is described by a frequency factor F_r which is equal to the product of the resonant frequency and the disk radius a; it is a constant for a given material; (iii) the mechanical Q, Q_m, of the circuit is independent of frequency and is a function of the material.

From (i), since C_0 is proportional to a^2 and to $1/t$ (where t is the thickness), C_1 is also proportional to a^2 and $1/t$. Also from (ii), since the frequency factor is given by $F_r = f_r a$, therefore $LC_1 = (a/2\pi F_r)^2$.

From this it follows that the thickness dependence of L is inverse to that of C_1, i.e. L is proportional to t.

Finally from (iii), and since $Q = \omega_r L/R$, R is proportional to thickness and inversely proportional to the radius. These results are summarized in Table 1.

Table 1. *Dependence of disk circuit parameters on dimensions*

Parameter	*Radius (a)*	*Thickness (t)*
Electrostatic capacitance C_0	a^2	t^{-1}
Mechanical capacitance C_1	a^2	t^{-1}
Coupling factor k	—	—
Mechanical inductance L	—	t
Mechanical resistance R	a^{-1}	t
Mechanical Q	—	—
Resonant frequency	a^{-1}	—
Frequency factor F_r	—	—

Second order thickness effect

It should be emphasized that the foregoing argument is somewhat idealized: it is found in practice that the resonant frequency is slightly dependent upon the radius/thickness ratio. Now, since this affects the ratio of the frequencies used to determine Poisson's ratio, the value of σ as calculated herein will also be thickness dependent. This is due to the neglect in the mechanical analysis of any perturbation of the boundary conditions due to finite thickness. Thus, for accurate assessment of a material, the values of the frequency ratio should be extrapolated to zero thickness.

The effect of non-zero thickness on Poisson's ratio and subsequently on coupling factor is shown in Figs 7 and 8.

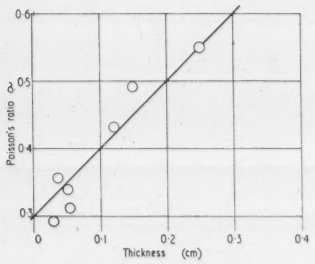


Fig. 7. Poisson's ratio σ plotted against thickness. Fundamental determination.

These values were determined experimentally using Fig. 2 and Fig. 4 respectively.

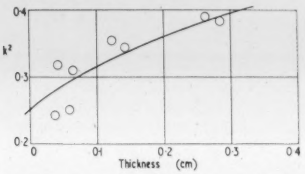


Fig. 8. k^2 plotted against thickness. Fundamental determination.

4. Modified electrical circuit

The most obvious shortcoming of the simple equivalent circuit is that it describes behaviour in the neighbourhood of only one resonance, so that a different circuit is required at each resonance.

The simplest modification is to add series resonant arms to allow for the overtone behaviour. If as is shown in Fig. 9

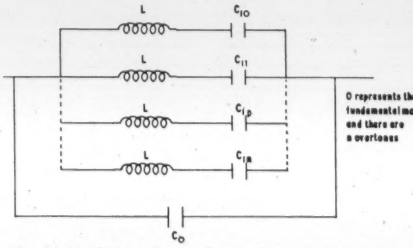


Fig. 9. Equivalent circuit of resonator including overtones.

the mechanical Q is assumed to be infinite, this leads to the following analysis.

The admittance of the fundamental arm is given by

$$Y_0 = \frac{j\omega C_{10}}{\omega^2 LC_{10} - 1}$$

so that the total admittance of the system will be

$$Y_T = \sum_{p=0}^{n} \frac{j\omega C_{1p}}{\omega^2 LC_{1p} - 1} + j\omega C_0 =$$

$$j\omega\left(\sum_{p=0}^{n} \frac{{\omega_{rp}}^2 C_{1p}}{\omega^2 - {\omega_{rp}}^2} + C_0\right). \quad (10)$$

Now, as the inductances are equal, one can write

$${\omega_{rp}}^2 C_{1p} = {\omega_{r0}}^2 C_{10}. \quad (11)$$

Substituting,

$$Y_T = j\omega C_{10}\left(\sum_{p=0}^{n} \frac{{\omega_{r0}}^2}{\omega^2 - {\omega_{rp}}^2} + \frac{C_0}{C_{10}}\right).$$

At any anti-resonance $\omega = \omega_{ap}$, $Y = 0$ and assuming that $\omega = \omega_{ap} = \omega_{rp}$ except at the anti-resonance concerned in which case $\omega_{ap} - \omega_{rp} = \Delta\omega_p$, then

$$\left(\frac{\omega_{r0}}{\omega_{rp}}\right)^2 \frac{\omega_{rp}}{2\Delta\omega_p} = \frac{C_0}{C_{10}} + \sum_{\substack{q=0 \\ p \neq q}}^{n} \frac{{\omega_{r0}}^2}{{\omega_{rq}}^2 - {\omega_{rp}}^2}. \quad (12)$$

Now, coupling factor has been defined as the ratio of stored mechanical energy to total stored energy, and applying this definition to the above circuit at frequency ω it can be shown that the coupling factor is given by

$$k^2 = \frac{\sum_{p=0}^{n} C_{1p}\left\{\frac{\omega^2\omega_{rp}^2}{\omega_{rp}^4+\omega^4}+\frac{\omega_{rp}^4}{(\omega_{rp}^2-\omega^2)^2}\right\}}{\sum_{p=0}^{n} C_{1p}\left\{\frac{\omega^2\omega_{rp}^2}{\omega_{rp}^4+\omega^4}+\frac{\omega_{rp}^4}{(\omega_{rp}^2-\omega^2)^2}\right\}+C_0}. \quad (13)$$

This equation can be evaluated to show the change in k^2 with frequency, and two cases are shown in Fig. 10 calculated

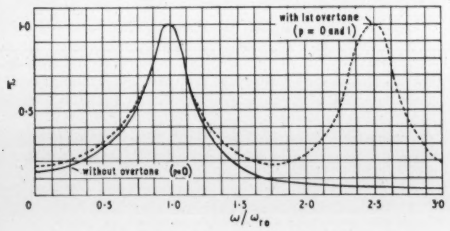


Fig. 10. Variation of coupling factor with frequency ($C_{10}/C_0 = 0{\cdot}2 : \omega_{r1} = 2{\cdot}5\ \omega_{r0}$).

for $p = 0$ and for $p = 0$ and 1 (i.e. fundamental and first overtone only). The ratio of C_{10} to C_0 has been taken to be equal to $0{\cdot}2$ ($\omega_{a0} = 1{\cdot}1 \times \omega_{r0}$), it has also been assumed that $\omega_{r1} = 2{\cdot}5\ \omega_{r0}$. It can be seen that k^2 is frequency dependent.

However at zero frequency Eqn (13) reduces to

$$k^2 = \sum_{p=0}^{n} C_{1p} \Big/ \Big(\sum_{p=0}^{n} C_{1p} + C_0\Big)$$

which is the more usual expression for k^2.

Using Eqn (11), this becomes

$$k^2 = \Big(1+\sum_{p=1}^{n}\frac{\omega_{r0}^2}{\omega_{rp}^2}\Big) \Big/ \Big\{\Big(1+\sum_{p=1}^{n}\frac{\omega_{r0}^2}{\omega_{rp}^2}\Big)+\frac{C_0}{C_{10}}\Big\}. \quad (14)$$

In certain special cases it is possible to evaluate $\sum_{p=1}^{n} \omega_{r0}^2/\omega_{rp}^2$ in terms of known series. Thus for a longitudinally vibrating bar in which only odd harmonics are excited

$$1+\sum_{p=1}^{n}\frac{\omega_{r0}^2}{\omega_{rp}^2} = 1+\frac{1}{9}+\frac{1}{25}+\ldots = \frac{\pi^2}{8}.$$

It is possible to substitute in Eqn (14) for C_0/C_1 from Eqn (12), so that finally we have for the coupling factor at some overtone p

$$k^2 = \frac{\Big(1+\sum_{p=1}^{n}\frac{\omega_{r0}^2}{\omega_{rp}^2}\Big)\Big(\frac{\omega_{rp}}{\omega_{r0}}\Big)^2\frac{2\Delta\omega_p}{\omega_p}}{1+\frac{2\Delta\omega_p}{\omega_p}\Big(\frac{\omega_{rp}}{\omega_{r0}}\Big)^2\Big(1+\sum_{p=1}^{n}\frac{\omega_{r0}^2}{\omega_{rp}^2}-\sum_{\substack{q=0\\q\neq p}}^{n}\frac{\omega_{r0}^2}{\omega_{rq}^2-\omega_{rp}^2}\Big)}. \quad (15)$$

Eqn (15) is a general expression applicable to any system. In the case of radial motion of disks it is possible to evaluate the frequency ratio in terms of the Poisson's ratio of the disk material; the roots and the ratio of the roots are tabulated in Table 2 for $\sigma = 0{\cdot}35$ and since $R_1/R_0 = f_{r1}/f_{r0}$, etc., it is possible to evaluate Eqn (15).

Table 2. *Roots and ratio of roots for* $\sigma = 0{\cdot}35$

Root	*Value*	*Ratio* R_p/R_0
R_0	2·08	—
R_1	5·40	2·60
R_2	8·58	4·13
R_3	11·74	5·64
R_4	14·89	7·16
R_5	18·04	8·67
R_6	21·18	10·18
R_7	24·33	11·70

The values obtained at the fundamental are given in Eqn (16).

$$k^2 = \frac{2{\cdot}58\Delta f/f_{r0}}{1+1{\cdot}90\Delta f/f_{r0}}. \quad (16)$$

For this same value of σ, Eqn (5) gives

$$k^2 = \frac{2{\cdot}63\Delta f/f_{r0}}{1+2{\cdot}63\Delta f/f_{r0}}.$$

The value of k^2 obtained with Eqn (16) may be compared with those from the graphical method. Table 3 gives some typical results using materials of different composition.

Table 3. *Comparison between coupling factor calculated from equation* (16) *and by the graphical method*

Graphical method	*Eqn 16*	*Graphical method*	*Eqn 16*
0·16	0·17	0·37	0·37
0·32	0·31	0·32	0·33
0·24	0·25	0·38	0·38
0·25	0·24	0·41	0·40
0·31	0·31	0·52	0·49

It can be seen that reasonable agreement exists.

5. Discussion

It has been shown that although discrepancies exist in the formulae at present used to evaluate the electromechanical coupling factor of a piezoelectric disk, these discrepancies are due to approximations in both the mechanical and equivalent circuit analyses. If the approximations are removed from the mechanical analysis and the effect of overtones included in the equivalent circuit analysis, good agreement is obtained between the two approaches.

Acknowledgment

The authors wish to thank the Plessey Co. for permission to publish this paper.

References

BOGDANOV, S. V., and TIMONIN, A. M., 1957, *Bul. Acad. Sci. U.S.S.R.*, **21**, 339.

KATZ, H. W., 1959, Ed., *Solid State Magnetic and Dielectric Devices* (New York: John Wiley), 91.

MASON, W. P., 1948, *Phys. Rev.*, **74**, 1134.

VAN DYKE, K. S., 1925, *Phys. Rev.*, **25**, 895.

The copper staining of p–n–p alloyed junction transistor sections

by S. RATCLIFFE and J. E. HUGHES, Ph.D., D.I.C. Associated Electrical Industries (Woolwich) Limited, Research Laboratory, West Road, Templefields, Harlow, Essex

MS. received 21*st November* 1960

Abstract

A technique is described for the staining of germanium p–n–p alloyed junction transistors using chemical deposition of copper. It is shown that copper is deposited preferentially on p-type regions and, in so doing, reveals certain features associated with regrowth processes. Other structural effects which are revealed may be connected with strain induced in the semiconductor by thermal cycling.

In an earlier paper Hughes and Ratcliffe (1959) have described the use of metallographic techniques for the rapid mounting, sectioning and examination of p–n–p indium–germanium alloyed junction transistors. It has now been shown that controlled deposition of copper on p-type regions can be used to facilitate the examination of p–n junctions and that, in addition, certain other features are revealed such as regrowth structures and strain markings.

Deposition of copper on the surfaces of semiconducting materials to reveal features such as diffusion boundaries and grown junctions has previously been applied but the use of an electrolytic cell has been considered desirable when applying this technique to germanium (Bond and Suits 1956, R.C.A. Laboratories 1956, Bridges, Scaff and Shive 1958, Turner 1959, Glang 1960). In the case of p–n–p germanium devices the presence of the indium renders an external potential source unnecessary. This is particularly useful when it is desired to stain a large number of devices mounted together but without electrical continuity through the samples.

The preferred technique is to immerse the polished sections in a solution, held at room temperature, consisting of 1 gramme of copper nitrate dissolved in 20 cm^3 of CP4 etching solution (HF 15 cm^3, HNO_3 25 cm^3, acetic acid 15 cm^3, bromine 0·3 cm^3). The concentration of copper has been found not to be critical and wide variations in practice may be tolerated. Periodic replacement of the solution is necessary.

Figure 1 shows a batch of experimental alloyed transistors. The clear junction delineation may be noted, although it may be observed that black and white reproduction cannot reveal the full advantage of a brightly coloured zone for examination purposes.

A number of additional features have been observed in experimental devices. Fig. 2 shows regrowth structures in the p-type germanium which are revealed as straight lines, behind regions of planar junction geometry, or as curved or

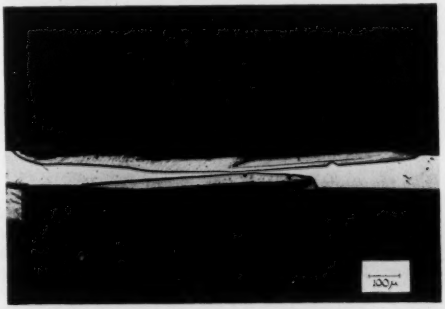


Fig. 2. Planar and curved regrowth markings revealed by copper staining.

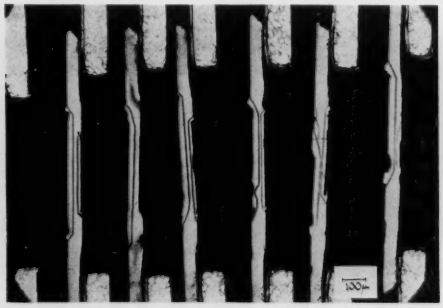


Fig. 1. Bulk mounted transistor sections with copper stained p regions.

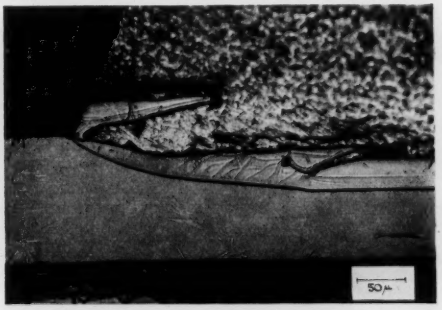


Fig. 3. Structural markings in germanium 'bridging' regions.

confused lines, behind regions where non-planar junctions occur. Fig. 3 illustrates structures which may be seen within the indium in the bridging regions, i.e. the non-epitaxial regrowth features which are propagated from nucleation points on the periphery of the indium. It is considered that the various structures observed are associated with alloying and regrowth processes and that useful information of the mechanisms operating during transistor manufacture may be obtained by their examination.

Further features have been observed in the regions of germanium adjacent to the wafer support (Fig. 4). These structures, which may arise from strain effects induced during the thermal cycling stages, are known to vary with the nature and composition of the support material and its solder coating.

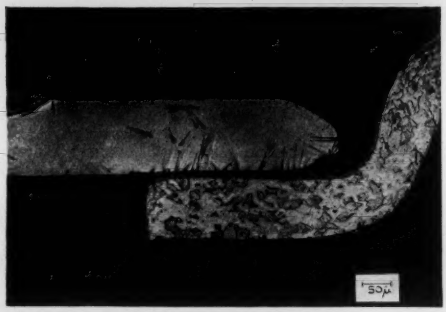


Fig. 4. 'Strain' markings near wafer support.

Acknowledgments

Thanks are due to R. Maddock for useful discussions and to Dr. M. E. Haine for permission to publish this note.

References

BOND, W. L., and SUITS, F. M., 1956, *Bell Syst. Tech. J.*, **35**, 1209.

BRIDGES, H. E., SCAFF, J. H., and SHIVE, J. N., 1958, *Transistor Technology*, **1**, 288 (Princeton, N.J.: Van Nostrand).

GLANG, R., 1960, *J. Electrochem. Soc.*, **107**, 356.

HUGHES, J. E., and RATCLIFFE, S., 1959, *Proc. Instn Elect. Engrs*, Pt. B, Suppl. No. 15, **106**, 451.

R.C.A. Laboratories, 1956, *Transistors*, **1**, 130.

TURNER, D. R., 1959, *J. Electrochem. Soc.*, **106**, 701.

CORRESPONDENCE

Comments on the A.S.T.M. Powder Data File—6

The writers have now prepared a sixth booklet, bearing the above title, which contains further additions to the lists previously published and described in the Institute's *Journals* (Hughes, Lewis and Wilson 1960). Copies of the present booklet have been distributed in the same way as were the others. Those who have not received copies of the booklets in this way may obtain them (free of charge) on request from the Institute and Society or from us.

It is hoped that similar lists of comments will continue to be published at approximately annual intervals. Users of the Data File are urged to send to the writers any material that they feel should be included in future issues.

It would help considerably in the work of extending the File if those who are working on structures by single-crystal methods could send us the crystallographic data and sufficient of their material to permit the production of a powder pattern from it. Full acknowledgment of the source would, of course, appear on the published card.

Viriamu Jones Laboratory,
University College, Cardiff,
Great Britain.

J. W. HUGHES
ISABEL E. LEWIS
A. J. C. WILSON

3rd January 1961

HUGHES, J. W., LEWIS, I. E., and WILSON, A. J. C., 1960, *Brit. J. Appl. Phys.*, **11**, 306; *J. Sci. Instrum.*, **37**, 183.

Notes and comments

Correspondence in *British Journal of Applied Physics*

The Editorial Board has considered the Correspondence Column of the *British Journal of Applied Physics* and hopes to expand this. We have therefore arranged with our printers to publish items of correspondence received up to six weeks before the date of publication, provided they are reported upon favourably by our referees. For example, items for publication in the **June** issue of the Journal should be received in the Office **not later than 20th April.**

Maxwell Colour Centenary

The Colour Group, in collaboration with The Institute of Physics and The Physical Society and the Inter-Society Colour Council of America, is holding a Conference on Tuesday, Wednesday and Thursday, 16th, 17th and 18th May 1961 at the Imperial College of Science and Technology, London, S.W.7, with a centenary discourse at the Royal Institution on 17th May.

This Conference is being arranged to mark the centenary of the famous demonstration of trichromatic colour reproduction given at the Royal Institution in London by James Clark Maxwell on 17th May 1861.

The subjects for discussion on the three days will be Trichromatic Principles, Colour Reproduction and Colour Appearance. Each session will be opened by a short review of the current position of the subject concerned, followed by papers by one or two well-known workers in the field; time will then be allowed for contributions by other conference members and for discussion.

Abstracts of the papers to be presented at the Conference will be published in the March/April issue of the *Journal of Photographic Science*.

Further details of the Conference are obtainable from The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7.

Engineering, Marine, Welding and Nuclear Energy Exhibition

The Engineering, Marine, Welding and Nuclear Energy Exhibition will be held at Olympia, London, from 20th April–4th May 1961. This biennial exhibition attracts close attention from buying authorities and technical personnel in every part of the world. The May issue of *Nuclear Engineering* will be a special issue containing a technical appraisal of equipment designed and manufactured for use in nuclear installations which will be exhibited at this Exhibition.

International Symposium on Microchemical Techniques

An International Symposium on Microchemical Techniques, organized by the Metropolitan Microchemical Society under the sponsorship of the International Union of Pure and Applied Chemistry, will be held at the Pennsylvania State University, U.S.A., from 13th–18th August 1961.

The technical programme will consist of a number of main introductory lectures and a full schedule of invited papers directed towards new methods and techniques or unique applications of microchemical or microanalytical interest.

A commercial exhibition of laboratory equipment, chemicals and specialities will be held in conjunction with the meeting.

Anyone interested in receiving further details should contact Mr. Howard J. Francis, Jr., Vice-Chairman of the International Symposium on Microchemical Techniques, c/o Pennsalt Chemicals Corporation, Post Office Box No. 4388, Philadelphia 18, Pennsylvania, U.S.A.

Radio and Electronic Component Show

The Radio and Electronic Component Show will be held this year in the Grand Hall at Olympia, London, instead of at Grosvenor House, where it has been situated for the past 16 years. The Show opens on Tuesday, 30th May until Friday, 2nd June inclusive.

International Congress Calendar

The *Annual International Congress Calendar* for 1961, produced by the Union of International Associations, is available from the Association at Palais d'Egmont, Bruxelles 1, Belgium, at a price of 18*s*. or 50*s*. in combination with one year's subscription to *International Associations*. This *Calendar* includes a chronological listing of international congresses, conferences, meetings and symposia scheduled to take place in 1961 and subsequent years. The following details are given: date, place, address of organizing body, theme, estimated number of participants, plans for publishing Reports or Proceedings, exhibition held in association with the meeting, and subject and geographical indexes. Supplements to the *Calendar* appear in the monthly magazine *International Associations*.

International School of Physics 'Enrico Fermi'

The following Summer Courses, organized by the International School of Physics 'Enrico Fermi' will be held this year at Varenna, Italy:

(1) 23rd May–3rd June	Cosmic rays, solar particles and space research.
(2) 19th June–1st July	Evidence for gravitational theories.
(3) 3rd July–15th July	Liquid Helium.
(4) 17th July–5th August	Semiconductors.
(5) 7th August–26th August	Nuclear Physics.

Applications should be received by the organizers not later than 25 days before the beginning of the Course. All correspondence concerning these Courses should be addressed to the Organizer and Director in charge of each Course, as follows:

(1) Professor B. Peters, Universitetets Institut for Teoretisk Fysik, Blegdamsvej, 17, Copenhagen, Denmark.
(2) Professor C. Møller, Universitetets Institut for Teoretisk Fysik, Blegdamsvej, 17, Copenhagen, Denmark.
(3) Professor G. Careri, Istituto di Fiscia dell'Universita, P. de delle Scienze, 5, Roma, Italy.
(4) Dr. C. A. Hogarth, Department of Physics, Brunel College of Science and Technology, Woodlands Avenue, Acton, London, W.3.
(5) Professor V. F. Weisskopf, Department of Physics, Massachusetts Institute of Technology, Cambridge 39, Mass., U.S.A.

United Nations Conference on New Sources of Energy

The United Nations Conference on New Sources of Energy, which will examine practical problems and experience in the utilization of solar energy, wind power and geothermal energy, with especial reference to the problems of the less developed countries, will be held in Italy (probably in Rome) from 21st–31st August 1961.

The Conference aims at bringing together experts in the fields of the application of solar energy, wind power and geothermal energy, as well as those interested in energy development in general, to provide participants with up to date information on progress achieved, and to facilitate an exchange of views and experience relating to practical problems in utilizing these sources of energy. The Conference will focus attention on applications rather than on the discussion of scientific principles and basic research. Theoretical studies will be discussed only if they appear to be closely related to practical developments. The subject of costs will be emphasized and stress will be laid on the various needs for energy and on specific means for meeting those needs.

Attendance at the Conference will be by invitation only.

The Proceedings and a Report of the Conference will be published later, as United Nations Sales publications.

Requests for further information should be addressed to the Executive Secretary of the Conference, United Nations, New York.

Mesucora 1961

Mesucora, the International Exhibition devoted to Measurement, Control, Regulation and Automation, will be held at the Exhibition Hall of the Centre National des Industries et des Techniques, Paris, from 9th–17th May 1961. It will cover an exhibition area of 35 000 square metres and bring together constructors from 13 European countries and the United States of America.

At the same time as the Exhibition, an International Congress will be held in Paris, the theme of which will be 'Recent progress in the fields of Measurement, Control, Regulation and Automation, resulting from the co-operation of the various techniques—electric, electronic, mechanical and nuclear.'

For further information apply to Mesucora, Service de Propagande, 40, rue du Colisée, Paris (8e).

Conférence Internationale des Arts Chimiques

The Sixth Salon International de la Chimie and Conférence Internationale des Arts Chimiques will be held from the 25th April–4th May 1962 at the Centre National des Industries et des Techniques, Rond-Point de la Défense, Puteaux-Paris.

For further details of the programme apply to the General Secretariat, Conférence Internationale des Arts Chimiques, 28 rue Saint-Dominique, Paris (7e).

Combustion Engineering Association Conference

A report on the Combustion Engineering Association's Conference on 'Getting the best out of Coal', held in Harrogate in November 1960, is now available and can be obtained from the Association at 70 Jermyn Street, London, S.W.1, price £1.

The report includes papers by leading authorities in their several subjects which examine from both the technical and economic aspects, the choice of coal allied to the selection, operation and maintenance of equipment for boiler plants and industrial furnaces and kilns, in order to achieve the maximum savings in capital and operating costs. Time was allowed for discussion at the Conference and these discussions are fully reported.

Industrial Finishes Exhibition

The first International Industrial Finishes Exhibition will be held at Earl's Court, London, from 8th–11th May 1961.

Journal of Scientific Instruments

Contents of the April issue

PAPERS

The high-precision isotopic analysis of uranium hexafluoride. By J. Bishop, D. F. Davidson, P. B. F. Evans, A. N. Hamer, J. A. McKnight, and E. J. Robbins.
A successful version of the Beevers–Macewan Fourier synthesizer using dekatron counters. By J. W. Jeffery.
Use of 'air-bleed' when measuring pressure in vacuum processes in which condensable vapours are present. By H. Wycliffe.
Correction term for dielectric measurements with cavity resonators. By E. S. Hotston.
Vacuum sluice valve of large aperture. By P. H. Lissberger.
Simple millimicrosecond pulse generator. By H. S. Caplan and D. T. Stewart.
A versatile metal to dielectric seal. By P. C. Conder, H. Foster, and H. A. H. Boot.
Low-inductance capacitor banks and linear pinched discharges. By A. H. Gabriel, V. T. S. Howell, E. Thornton, and R. J. Wilson.
Calorimeter for the continuous study of the heat production of microbial systems. By W. W. Forrest.
A recording flowmeter using a radioactive float. By D. W. Tims.
Electrolytic tank for studying the flow of liquids through tubes of various cross sections. By M. J. Denton and S. G. Lipson.
X-ray diffraction data from single crystals at 20° K. By J. D. Forrester.
Technique using ^{85}Kr-labelled krypton for surface area measurements. By D. W. Aylmore and W. B. Jepson.

RESEARCH NOTES

A temperature compensated magnetizing coil. By J. E. Knowles.
Liquid nitrogen cryostat for electrical measurements on crystal samples in magnetic and light fluxes. By C. M. Hurd.

LABORATORY AND WORKSHOP NOTES

A level controller for liquid nitrogen. By D. Miller and H. C. Evans.
Power oscillator for small synchronous motors. By D. Allenden.
A simple plasma oscillator. By G. C. McCullach.
The production of high-quality optical surfaces on calcite. R. N. Smartt.
Mercury switch for establishing the true electrical zero in precise thermocouple measurements. By J. Middlehurst.
Cold hearth zone refining. By R. M. Ware.
An improvement to the floating zone method of growing single crystals. By G. W. Green.

CORRESPONDENCE

Comments on the A.S.T.M. Powder Data File—6. By J. W. Hughes, Isabel E. Lewis, and A. J. C. Wilson.

NOTES AND NEWS

New instruments, materials and tools. New books. Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.

REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611–2.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12s. 6d. each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.

CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

UNWIN BROTHERS LIMITED, WOKING AND LONDON

Amplification through stimulated emission—the Maser

by R. A. SMITH, Royal Radar Establishment, Great Malvern, Worcs.

MS. received 5th January 1961

Abstract

The various factors which determine the ultimate sensitivity of a radio receiver are considered. External sources of noise (random electromagnetic fluctuations) are examined and the conditions determined under which ultimate sensitivity is limited by noise generated in the receiver itself.

The physical principles of emission and absorption of radiation are examined with particular reference to their application to the microwave region of the electromagnetic spectrum. The use of stimulated emission for amplification (Maser action) is considered and various methods for its exploitation are discussed. Different types of Maser amplifier are described and an elementary method of calculating their noise factors is given.

Recent applications of Maser amplifiers to radio astronomy and communications are described and some figures given to illustrate the large gain in ultimate sensitivity which has been obtained. Thus it is now possible to design an amplifier with noise equivalent to the input from a load at a temperature of only a few degrees K, *and a whole system, including antenna and external noise from outer space, with noise equivalent to the input from a load at about* 20° K.

1. Limitation of sensitivity by noise

The smallest radio signal which can be detected and intelligibly displayed by means of a well-designed receiver is determined by the background noise which inevitably accompanies the signal. The signal will cease to be intelligible when it is so small that it is swamped by the background noise. Unless the gain of the receiver is sufficiently high to make this background noise observable it is inadequately designed for minimum signal reception. The amount of noise power appearing at the output naturally depends on the receiver gain, but then so does the strength of the desired signal and in a well-designed receiver the ratio of signal to noise should not depend on the gain. The amount of noise also depends on the bandwidth of the receiver, i.e. on the range of frequencies Δf over which it is sensitive for a fixed setting of the tuning. The bandwidth Δf can be given a precise numerical meaning,* but for the moment we shall not define it more exactly. Except when it is very large the noise power is found to be proportional to Δf and we may speak of the noise power per unit bandwidth—a quantity having the dimensions of energy. Another quantity related to the bandwidth is the response time τ of the receiving system; this again can be given a precise numerical meaning,* but may be thought of as the smallest time in which significant variations of signal are observable. It is related to the bandwidth by means of the equation $\tau = 1/4\Delta f$. The noise power per unit bandwidth is therefore approximately equal to the noise energy received in a time 4τ.

For maximum sensitivity, the bandwidth used will, in general, be the minimum value compatible with reception of the information in the desired signal, i.e. it will depend on the frequency spread of the signal. For example, for intelligible speech a bandwidth of 3 kc/s is adequate, while 'high-fi' enthusiasts demand a bandwidth of at least 15 kc/s, and for radar with 1 μs pulses a bandwidth of the order of 1 Mc/s is required. It would therefore appear necessary to specify, in each case, the bandwidth, in order to describe how good a receiver is. This is rather a nuisance and we shall see later how to overcome the difficulty. We must be quite clear, however, that in order to determine the signal-to-noise ratio obtained with a particular signal the bandwidth must be specified.

2. Sources of noise

Let us now examine briefly the main sources of noise which contribute to the total background noise in a radio receiver. They are as follows:

(*a*) locally generated electrical noise, such as that due to car ignition, etc.,
(*b*) electrical noise of terrestrial origin such as that due to thunderstorms—a form of noise often referred to as 'atmospherics',
(*c*) radio noise from outer space and from the sun,
(*d*) thermal noise generated by external sources and in the aerial system,
(*e*) noise generated in the receiver itself.

Source (*a*) can be troublesome at all frequencies, but since it is not fundamental in character, we shall not discuss it here. Most of the noise in the very low frequency radio bands comes from source (*b*). In the high frequency and very high frequency bands, at least down to wavelengths of the order of 1 m, the main source is (*c*). This amounts to saying that for wavelengths longer than about 1 m conventional radio receivers generate so little noise themselves that this is negligible compared with external sources of noise and there is no point in trying further to reduce receiver noise. For shorter wavelengths, particularly in the microwave region of the spectrum, noise generated in the receiver itself predominates, and anything that can be done to reduce this noise (keeping the gain constant) will lead to an improvement of signal-to-noise ratio. The main sources of noise in a receiver are†: e_1, thermal fluctuations in the input circuit; e_2, shot noise, current noise or flicker noise in the first valve or crystal.

The contribution e_1 is fundamental and depends only on

* See, for example, R. A. Smith, F. E. Jones and R. P. Chasmar 1957, pp. 231–244.

† A full discussion of these and their relative importance has been given in a number of books; see, for example, R. A. Smith, F. E. Jones and R. P. Chasmar 1957, pp. 178–204.

the absolute temperature T, the available noise power being equal to kT per unit bandwidth, where k is Boltzmann's constant.* Only for extremely sensitive microwave receivers is the contribution (d) important, but for some modern receivers this source of noise may predominate.

3. Noise factor and noise temperature

We must now see how to specify the quality of a receiver. Suppose first of all that the receiver is connected to a matched resistive load at temperature T_s (see Fig. 1). The input noise power per unit bandwidth P_{n1} will be equal to kT_s. If G is the power gain of the receiver the output noise power P_{n2} would be equal to GP_{n1} provided the receiver itself contributed no noise. In general, it will contribute noise power of amount P_a per unit bandwidth and we may write for the output noise power P_{n2}

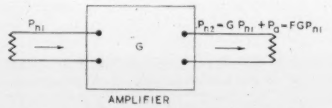


Fig. 1. Amplifier input and output noise.

$$P_{n2} = GP_{n1} + P_a. \tag{1}$$

The noise factor F of the receiver may then be defined by means of the equation

$$F = P_{n2}/GP_{n1} = 1 + P_a/GP_{n1}. \tag{2}$$

For a noiseless receiver $F = 1$. The quantity F as defined above depends on the noise power generated by the input load; the latter is therefore usually taken to be a resistive load at 300° K. Although the quantity F specified in this way is widely used, an alternative procedure is becoming more common. For any input we may *define* an effective temperature T_s so that $P_{n1} = kT_s$. For example, a lossless matched aerial looking at a uniform 'black' background at temperature T would have a noise temperature T. For any background the effective temperature T_e may be defined in the same way, for a given frequency, the input noise power being kT_e. When the aerial is not lossless, and itself contributes thermal noise, we may define an effective input temperature T_s which combines these sources to give a noise power per unit bandwidth kT_s. In the same way we may define a noise temperature T_r of a receiver by writing

$$P_a = GkT_r. \tag{3}$$

In this way we may express the noise factor F in the form

$$F = 1 + T_r/T_s. \tag{4}$$

In the above discussion we have added the noise powers from the various sources since they are incoherent. In practice this is not always strictly permissible, since the value of the input load may affect the characteristics of the input circuits, so that P_a may depend on P_{n1}; for simplicity we shall, however, assume that this procedure is permissible. The quantity T_r may then be used to specify the quality of the receiver, an ideal receiver having $T_r = 0$. The quantity F is sometimes expressed as a power ratio in decibels; the noise temperature is usually expressed in °K. Conventional microwave receivers have a noise figure not better than about 6 dB or $F = 4$, a value of 9 dB or $F = 8$ being more common. The corresponding receiver noise temperatures would therefore be 900° K, and 2100° K respectively.

The advantages of noise temperature as a criterion of receiver performance will be more evident when we later consider receivers whose noise temperature is only a few degrees Kelvin in which case the quantity F is very nearly equal to 1 and rather insensitive to variations in receiver noise temperature. Such receivers have now become a practical possibility with the invention of the *Maser*, the basic principles of which we now discuss.

4. The Boltzmann distribution

Later, we must consider in some detail how radiation is emitted and absorbed by an isolated atom or molecule. When we say 'isolated' we do not mean *completely* isolated, but interacting with its neighbours so as to be in thermal equilibrium, yet sufficiently weakly so as not to have its energy levels appreciably disturbed. Let us consider only two energy levels for simplicity and suppose that almost all the molecules have either energy E_1 or E_2 ($E_2 > E_1$). These are generally sharply defined quantum levels and the molecule can have no other energies near these. An atom having energy E_1 (or as we say in energy level 1) can absorb radiation by being excited up to energy E_2, the frequency ν_{12} of the absorbed radiation being given by

$$h\nu_{12} = E_2 - E_1. \tag{5}$$

To this approximation this is a 'line' absorption, other frequencies not being absorbed. The molecule may also emit radiation by jumping from level 2 to level 1, the frequency of the emitted radiation being also ν_{12} as given by (5).

Now suppose we have a large number N of molecules, then the ratio of the number N_1 in level 1 to the number N_2 in level 2, in thermal equilibrium, is given by the well-known Boltzmann law

$$N_2/N_1 = \exp\{-(E_2 - E_1)/kT\} = \exp\{(-h\nu_{12}/kT)\}. \tag{6}$$

Let us now consider some numerical values. The value of k is $1{\cdot}38 \times 10^{-16}$ erg/deg K and a useful value to remember is that kT expressed in electron volts (kT is an energy) is about 1/40 ev when $T = 300°$ K (nominal room temperature). Another order of magnitude worth remembering is that $kT \simeq 10^{-4}$ ev when $T = 1°$ K. Now, some values of $h\nu$ worth remembering are $h\nu \simeq 1$ ev (actually $1{\cdot}2$) when the wavelength $\lambda = c/\nu$ is 1 μ, i.e. in the near infra-red. In the middle of the visible spectrum (green) $\lambda = 0{\cdot}5\ \mu$ and $h\nu \simeq 2$ ev, and for $\lambda = 1$ cm, $h\nu \simeq 10^{-4}$ ev. Thus for visible light $h\nu/kT \simeq 80$ when $T = 300°$ K. From the above values we see that if E_1 and E_2 are separated by about 2 ev so that emission is in the visible spectrum then $N_2/N_1 \simeq e^{-80}$ when $T = 300°$ K, i.e. $N_2 \ll N_1$ in thermal equilibrium. In considering absorption of radiation we therefore *need only take account of the molecules in the lower state.* If, however, we are dealing with microwave radiation, with say $\lambda = 10$ cm, then $h\nu/kT \simeq 4 \times 10^{-4}$, so that $h\nu/kT \ll 1$ and we have $N_1 \simeq N_2$. We may then write Eqn (6) in a more convenient form as follows:

$$\frac{N_1 - N_2}{N_1 + N_2} = \frac{e^{h\nu/kT} - 1}{e^{h\nu/kT} + 1}.$$

Expanding the exponentials we have

$$e^{h\nu/kT} \simeq 1 + h\nu/kT$$

* This is only so when the frequency ν is not too high and the temperature not too low. More generally, the noise power per unit bandwidth is equal to $h\nu/\{\exp(h\nu/kT) - 1\}$ where h is Planck's constant; this reduces to kT when $h\nu \ll kT$.

and since $N_1 + N_2 = N$

$$\frac{N_1 - N_2}{N} \simeq \frac{h\nu}{2kT}. \tag{7}$$

Thus, for $\lambda = 10$ cm and $T = 300°$ K, $N_1 - N_2 = 2 \times 10^{-4} N$. Even when $T = 4°$ K (temperature of liquid helium boiling under atmospheric pressure) we have

$$N_2 - N_1 \simeq N/80.$$

For $T = 1°$ K and $\lambda = 1$ cm, $h\nu \simeq kT$ and $N_2 = N_1/e$. Thus for most microwave applications we may assume that $h\nu \ll kT$ and use Eqn (7) and also kT for the noise power per unit bandwidth (see § 2).

5. Emission and absorption of radiation

Let us now consider how the populations (as they are called) N_1 and N_2 are affected by emission and absorption of radiation. There are two obvious processes by which the value of N_2 might be changed. The first is spontaneous emission of radiation. This is a random process and so depends only on the number of molecules N_2 and is proportional to this number. For this process we may therefore write

$$\left(\frac{dN_2}{dt}\right)_{\text{spont. em.}} = -AN_2 \tag{8}$$

where A is constant. The second obvious process is excitation from level 1 to level 2. The number excited is clearly proportional to N_1 and to the intensity of the radiation field in which the molecule is placed. Thus we have

$$\left(\frac{dN_2}{dt}\right)_{\text{excit.}} = BU(\nu)N_1 \tag{9}$$

where $U(\nu)d\nu$ is the energy density of the radiation field in the frequency interval $d\nu$, and B is a constant.

These two processes are, however, not the only ones which operate. As first shown by Einstein a third must be introduced known as 'stimulated emission'. Einstein showed that in the presence of a radiation field the probability of a transition from the *upper* level to the lower is increased by an amount proportional to the intensity of the field. Therefore we must write

$$\left(\frac{dN_2}{dt}\right)_{\text{stim. em.}} = -CU(\nu)N_2. \tag{10}$$

We shall now show how the constants A, B, C may be determined. If stimulated emission did not take place we should find that $C = 0$. In fact what we shall find is that $B = C$ and that the transition probability for stimulated emission is just equal to that for excitation (absorption). Combining (8), (9) and (10) we have

$$dN_2/dt = -AN_2 + BU(\nu)N_1 - CU(\nu)N_2. \tag{11}$$

In thermal equilibrium we must have $dN_2/dt = 0$. If we write $U_0(\nu)$ for the radiation field in thermal equilibrium we have then, on equating to zero the right-hand side of Eqn (11)

$$U_0(\nu) = \frac{A}{BN_1/N_2 - C}. \tag{12}$$

In thermal equilibrium N_1/N_2 is given by Boltzmann's equation (Eqn (6)) so that we have

$$U_0(\nu) = \frac{A}{Be^{h\nu/kT} - C}. \tag{13}$$

But $U_0(\nu)$ is well known and is given by Planck's formula*

$$U_0(\nu) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/kT} - 1}. \tag{14}$$

On comparing Eqn (13) with Eqn (14) we see at once that we must have $B = C$. There are thus only two coefficients, A, B, required, and these are known as the Einstein coefficients. *This is one of the fundamental results of radiation theory and on it depends the whole of the working of molecular amplifiers and oscillators.*

From Eqns (13) and (14) we get a second important result; we have now

$$\frac{A}{B} = \frac{8\pi h\nu^3}{c^3}. \tag{15}$$

In thermal equilibrium the ratio of the probability of a spontaneous emission to that of a stimulated emission, as will be seen by comparing Eqns (8) and (10), is given by

$$\frac{A}{BU_0(\nu)} = e^{h\nu/kT} - 1. \tag{16}$$

Thus, in the visible region of the spectrum, this ratio is very large indeed and except in very intense radiation fields and at high temperatures spontaneous emission is vastly more probable than stimulated emission. In the microwave region of the spectrum, however, since generally $h\nu \ll kT$ we have

$$\frac{A}{BU_0(\nu)} \simeq \frac{h\nu}{kT} \ll 1 \tag{17}$$

and spontaneous emission is generally much less probable than stimulated emission. Equation (17) refers to conditions of thermal equilibrium, and if we try to apply it to compare spontaneous emission with that produced by monochromatic radiation of frequency ν_{12} we find ourselves in a slight difficulty. The energy density in strictly monochromatic radiation is infinite and we should then conclude that stimulated emission is infinitely more probable than spontaneous emission. This is, of course, incorrect. The difficulty is resolved when we appreciate that the absorption line corresponding to the transition has finite width $\Delta\nu$ arising from a number of causes, some of which we shall discuss later. We may therefore compare the incident monochromatic power $P(\nu)$ with the effective power in the equilibrium radiation field $P_0(\nu)\Delta\nu$. In an experiment this will generally just be the noise power in a bandwidth corresponding to the line width and if we use power $P(\nu)$ greater than this (as we generally shall) then *a fortiori* the stimulated emission will predominate. The result is that for most practical purposes we may ignore *spontaneous* emission in the microwave region of the spectrum.

6. Emission from a discharge

It is instructive to consider the emission of radiation from an electric discharge in terms of these ideas and to understand why the process of excitation which works so well at optical frequencies is so very inefficient for microwaves. It is also instructive to consider the nature of the emitted radiation. Suppose we have a steady-current discharge; electrons are accelerated in the electric field and attain energies capable of exciting optical levels by electron impact. Once excited, the atoms may emit the energy of excitation or

* See, for example, R. A. Smith, F. E. Jones and R. P. Chasmar 1957, p. 29, Eqn (29); this must be divided by $c/4$ to get the energy density from the power flux, cf. p. 23, Eqn (2).

perhaps pass it on in the form of kinetic energy in what are called 'collisions of the second kind'. Because of the high probability of spontaneous emission (short lifetime τ_s) the emission process is the one most likely to take place and an appreciable amount of the energy given to the electrons reappears as light emitted in this way. Since spontaneous emission is a random process the emitted radiation will be incoherent and will be in the form of narrow-band noise. Now in the microwave region of the spectrum the probability of spontaneous emission is so small that this process cannot compete with other ways of taking away the energy of excitation, e.g. collisions with the walls. (Microwave energy is in fact emitted in the form of wide-band noise but this comes mainly from the free electrons in the discharge and is not a characteristic radiation.) Because of the relatively high probability of stimulated emission in this region of the spectrum it is to this that we must look for a possible emission process. Before considering this we must examine the absorption process in rather more detail.

7. Absorption of microwave radiation

In view of the above discussion we shall, in what follows, neglect spontaneous emission. In considering absorption at optical frequencies, except in very intense radiation fields at high temperatures (e.g. in stars) we may neglect the effect of stimulated emission on the absorption of the radiation. For microwaves this is certainly not so and the stimulated emission has just as great an effect in determining the amount of radiation absorbed as the excitation process. Neglecting then spontaneous emission we may write Eqn (11) in the form

$$\frac{dN_2}{dt} = -BU(\nu)(N_2 - N_1) \tag{18}$$

also since

$$\frac{dN_2}{dt} = -\frac{dN_1}{dt}$$

$$\frac{dN_1}{dt} = -BU(\nu)(N_1 - N_2). \tag{19}$$

Now suppose we have a power P at frequency ν in a beam of unit cross-sectional area, so that the number N_ν of quanta incident on unit area per unit time is $P/h\nu$. Now consider a small thickness dx; the power absorbed per unit time is just $h\nu(dN_1/dt)dx$. We may write Eqn (19) in terms of P instead of $U(\nu)$ which is proportional to it, i.e. we have

$$\frac{dN_1}{dt} = -aP(N_1 - N_2) \tag{20}$$

where a is a transition probability per unit incident power. Thus we have

$$\frac{dP}{dx} = h\nu\frac{dN_1}{dt} = -h\nu aP(N_1 - N_2). \tag{21}$$

The absorption coefficient α is defined by the equation

$$\frac{dP}{dx} = -\alpha P \tag{22}$$

so that

$$P = P_0 e^{-\alpha x}. \tag{23}$$

Hence we see that α is given by

$$\alpha = h\nu a(N_1 - N_2). \tag{24}$$

The absorption is therefore proportional to the *difference* between the populations of the two levels. When $h\nu \ll kT$ we then have from Eqn (7)

$$\alpha = ah^2\nu^2 N/2kT,$$

and we see that α increases as T is lowered.

Now it is well known that the process of *absorption* is a coherent one, i.e. as a wave with well-defined phase passes through an absorbing medium its phase may *change*, but the coherence is not destroyed. It follows that stimulated emission is also a coherent process as (24) shows that the net absorption is just a balance between direct absorption and stimulated emission. Thus if we could obtain more stimulated emission than direct absorption we should have net emission and the radiation so emitted would be coherent provided the exciting radiation is coherent; such a process can be truly described as 'amplification'. When we examine Eqn (21) we see that the condition for net emission ($dN_1/dt>0$) is that $N_2 > N_1$. This can never occur in thermodynamic equilibrium and we must now consider methods of upsetting the thermodynamic equilibrium in order to make $N_2 > N_1$. The technique of designing molecular amplifiers consists of devising methods of achieving this condition, and the rest of this article will be mainly concerned with a description of the various methods which have been used. Before proceeding with this, however, we must consider the effect of absorption in changing the ratio N_2/N_1.

8. Saturation

When radiation is absorbed as described in the last section it is quite clear that the populations in the two levels will begin to change. If at $t = 0$ we have thermodynamic equilibrium so that $N_1 > N_2$ then $dN_1/dt < 0$ and $dN_2/dt > 0$, i.e. the population of the lower level begins to decrease and that of the upper level to increase. If there were no processes tending to restore the equilibrium condition this would clearly continue till $N_1 = N_2$ when no more radiation would be absorbed; this condition is called 'saturation'. We may readily determine the time-constant with which it is reached from Eqn (20). Inserting $N_2 = N - N_1$ in (20) we have

$$\frac{dN_1}{dt} = -aP(2N_1 - N). \tag{25}$$

Equilibrium is clearly established when $N_1 = N_2 = \frac{1}{2}N$ for then only is $dN_1/dt = 0$. The solution of (25) is of the form

$$N_1 = De^{-2aPt} + \tfrac{1}{2}N. \tag{26}$$

We determine the constant D so that at $t = 0$ $N_1 = N_{10}$ giving

$$N_1 = (N_{10} - \tfrac{1}{2}N)e^{-2aPt} + \tfrac{1}{2}N. \tag{27}$$

The time constant τ for saturation is equal to $(2aP)^{-1}$. High power and large transition probability give a short time to achieve saturation.

In the above discussion we have omitted to consider processes such as collisions, etc., tending to restore thermodynamic equilibrium. These may generally be represented as having a relaxation time τ_0, i.e. if the system were left to itself it would decay exponentially back to equilibrium with time-constant τ_0, i.e. we should have

$$\frac{dN_1}{dt} = -\frac{N_1 - N_{10}}{\tau_0}. \tag{28}$$

Adding this process to that due to absorption we have, combining Eqns (28) and (25),

$$\frac{dN_1}{dt} = -aP(2N_1 - N) - \frac{N_1 - N_{10}}{\tau_0}$$
$$= -\left(2aP + \frac{1}{\tau_0}\right)N_1 + aPN + \frac{N_{10}}{\tau_0}. \qquad (29)$$

Thus we see that the time constant τ for equilibrium is now given by

$$\frac{1}{\tau} = \frac{1}{\tau_0} + 2aP$$

and the equilibrium values N_{1s} and N_{2s} are

$$N_{1s} = \frac{aPN + N_{10}/\tau_0}{2aP + 1/\tau_0} \qquad (30)$$

$$N_{2s} = \frac{aPN + N_{20}/\tau_0}{2aP + 1/\tau_0}. \qquad (30a)$$

If $\tau_0 \ll 2aP$ we have $N_{1s} \simeq N_{10}$, $N_{2s} \simeq N_{20}$, and the original equilibrium is hardly disturbed. If, on the other hand, $\tau_0 \gg 2\,aP$, $N_{1s} \simeq N_{2s} \simeq \frac{1}{2}N$. The power required to produce marked saturation is given by the condition

$$P \gg \frac{1}{2a\tau_0}.$$

A small transition probability and short relaxation time mean that a large power will be required. This process of saturation plays a central part in molecular amplifier technique.

By the above process we may effectively remove the absorption, but have not yet succeeded in getting an excess of stimulated emission. We now proceed to discuss the various methods by means of which this may be achieved, all of which depend on devising a means whereby N_2 may be made greater than N_1. This process, termed **M**icrowave **A**mplification through **S**timulated **E**mission of **R**adiation, was proposed independently by Gordon, Zeiger and Townes (1955) and by Basov and Prokhov (1956); devices based on its use are termed *Masers*.

9. Atoms or molecules in a solid

As some of the devices we shall consider use atoms or molecules in a solid to give stimulated emission we must digress for a moment to consider the effect of crowding of these together in the solid state. We specifically stated before that we were considering nearly isolated molecules and this approximation holds very well for a gas except at high pressures. For a solid two main effects take place which we have not so far considered. These are as follows:

(i) An excited molecule may fairly readily give up its energy of excitation to the rest of the solid in the form of lattice vibrations. (This is not a true spontaneous emission, but the effect is sometimes regarded as such and even called spontaneous emission in the literature. This we think is confusing and prefer to call it 'relaxation' since the process is one tending to restore thermal equilibrium.) It is found that this process may generally be represented by a relaxation time as in § 8. The relaxation time is generally called τ_1 in the literature and may vary from 10^{-8} s to several seconds When the energy levels 1 and 2 are derived from a spin system it is generally called the spin–lattice relaxation time.

(ii) Because of their mutual interactions the energy levels are broadened into narrow bands. (This always happens with a 'many-particle' situation, cf. the energy bands for electrons in a solid.) In order to keep the width of these bands down, 'dilute' materials are used, the active molecules being widely separated in the solid. The absorption lines are broadened and have a line width $\Delta\nu$, and we may define a time τ_2 by means of the relationship $\Delta\nu = 1/2\pi\tau_2$; τ_2 is sometimes also referred to as a relaxation time. For spin systems it corresponds to spin–spin interactions and is generally called the spin–spin relaxation time. In many cases, however, it is not a true relaxation time as τ_1 is.

10. The separation-type *Maser*

In principle, the simplest way of arranging that the number of molecules in level 2 exceeds the number in level 1 is to remove those in level 1 entirely. Suppose the state of excitation is used to 'label' the molecules in state 2 and the mixture passed through a chamber in which those labelled (2) are passed and those labelled (1) rejected. We should then be able to pass a stream of excited molecules into a waveguide or resonant cavity where they could be induced to emit by stimulation. A resonant cavity is generally preferred in view of the higher fields obtained with small incident power. The idea of separation is by no means so far-fetched as may seem at first sight and was used to make the first successful *Maser* by C. H. Townes and his staff at Columbia University (Gordon, Zeiger and Townes 1955).

The ammonia molecule exhibits very strong microwave absorption at a number of frequencies near 23·9 kMc/s (Gc/s). The main absorption is due to the so-called inversion of the molecule, fine structure being due to rotation. For simplicity, we shall only consider two levels. The form of the ammonia molecule NH_3 is illustrated in Fig. 2. The nitrogen atom

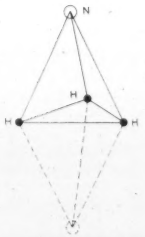


Fig. 2. The ammonia molecule.

has two positions of equilibrium, one above and one below the plane of the three hydrogen atoms. By the quantum-mechanical tunnel effect it can pass from one to the other. It is a well-known result of quantum mechanics that the apparently degenerate state corresponding to the two positions is split into two separated by energy ΔE. The frequency ν given by $\nu = \Delta E/h$ corresponds in a semi-classical picture to the frequency with which the nitrogen atom oscillates between the two positions, and is called the inversion frequency. The two levels we denote as before by (1) and (2). Were it not for this rapid inversion, the molecule would have a strong dipole moment because of its asymmetry, but because of the inversion this is a rapidly oscillating dipole and has no average moment. When an electric field is applied, however, it encourages the nitrogen atom to stay longer on one side of the plane of the hydrogen atoms and so *induces* a dipole moment. It turns out that because of the different symmetry of the wave functions for states (1)

and (2) the sign of the dipole moment is opposite for these states. For the lower state it is such as to decrease the energy in an electric field. If $\mathscr{E}$ is the strength of the electric field the induced dipole moment μ will be of the form $\mu = \gamma\mathscr{E}$. The energy of the molecule in state (1) will be given by

$$E = E_1 - \mu_1\mathscr{E} = E_1 - \gamma_1\mathscr{E}^2. \tag{31}$$

In state (2) it will be given by

$$E = E_2 - \gamma_2\mathscr{E}^2. \tag{32}$$

It turns out that $\gamma_1 > 0$ and $\gamma_2 < 0$. The levels therefore show a quadratic Stark effect, the electric field lowering the energy of the bottom level and raising that of the top. The different dipole moments may thus be used as 'labels'. Separation is achieved by use of an inhomogeneous electric field. It is clear that molecules in state (2) will tend to move into the weaker parts of the field and those in state (1) into the stronger parts. Ammonia molecules are shot from an oven through an electrostatic separator and the excited molecules are passed into a microwave cavity (Fig. 3). The

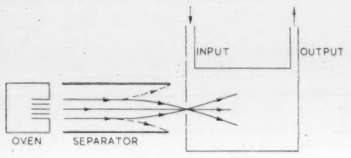


Fig. 3. Separation scheme for ammonia *Maser*.

separator consists of a cage of wires charged alternately positive and negative. Along the axis of the cage there is zero field (Fig. 4). The molecules in state (2) tend to move

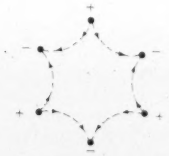


Fig. 4. Electrostatic separator.

towards the weak field near the axis and are focused, whereas those in state (1) are dispersed. If weak microwave power is fed in at the input to the cavity resonator, provided the amplification due to stimulated emission is greater than the loss in the cavity, net amplification will result. This will only occur if the cavity Q is high enough and it is properly loaded. If the cavity is lightly loaded, and the energy produced by stimulated emission exceeds that due to losses in the cavity, oscillation will take place. In this case the *Maser* will act as a very stable microwave oscillator, the frequency being essentially that of the ammonia absorption line being used.

If $\bar{P}$ is the average probability that a molecule will be de-excited in passing through the cavity the power generated by stimulated emission when N molecules pass through per second is equal to $N\bar{P}h\nu$. Now $\bar{P}$ is proportioned to the microwave energy W in the cavity. Let us therefore write $\bar{P} = CW$ where C is a constant determined by the transition probability, average time in cavity, etc. We then have for the power generated by stimulated emission $Nh\nu CW$. The power lost through cavity losses is given, as is well known, by $2\pi\nu W/Q$ (Q is the Q of the cavity). We have therefore oscillations if

$$Nh\nu CW > 2\pi\nu W/Q$$

or

$$N > 2\pi/hCQ. \tag{33}$$

There is thus, for a given value of Q, a critical number of molecules required to make the system oscillate.

As an amplifier this type of *Maser* has a certain bandwidth. The relaxation times discussed in § 9 do not apply here. Mutual interaction is negligible, since collisions between molecules, at the pressures used have little effect. There is a small spread in frequency due to Doppler shift but, since the cavity is designed to have its electric field at right angles to the line of flight of the molecules, the effect is very small. The effective relaxation time τ is in fact determined by collision with the walls of the cavity, and is nearly equal to the time of flight through the cavity. Classically, the molecule is only being stimulated for a finite time τ and radiates only during this time. This gives rise to a frequency spread of the order of $1/\tau$. The velocity V of the molecules from the oven as used is about 4×10^4 cm/s and the cavity has a length l of about 10 cm; thus $\tau = l/V = 2{\cdot}5 \times 10^{-4}$ s. The frequency bandwidth $\Delta\nu$ is thus only about 4 kc/s; this is one of the main limitations of the ammonia *Maser* as an amplifier. As an oscillator it appears to be extremely stable having a stability of the order of a few parts in 10^{12}, but the power generated is only of the order of 10^{-10} watt. As an amplifier, noise factors very close to 0 dB have been observed. Ammonia *Masers* have been made to operate in various laboratories and are generally used as highly stable oscillators or as microwave spectrometers. More recently a separation *Maser* of this type has been made to operate using atomic hydrogen giving an extremely narrow emission line at the frequency of the so-called 'galactic' hydrogen line (1420 Mc/s) (Goldenberg, Kleppner and Ramsey 1960). This promises to provide perhaps the most stable and reproducible oscillator yet devised and may have important uses as an 'atomic clock'.

11. The inversion-type *Maser*

If by some means or other we could reverse the positions of the levels (1) and (2) when in thermal equilibrium, without changing their populations, we should then have the greater number of molecules in the higher level and so get emission until relaxation restored the equilibrium distribution. Now suppose the two levels arise from a single unpaired electron in an atom. Such an electron may take up one of two positions in a magnetic field—its magnetic moment μ pointing either along the field or in the opposite direction. In the former case the magnetic energy is $-\mu H$ and in the latter it is μH. The two states arising in this way have energies

$$E_1 = E_0 - \mu H$$

$$E_2 = E_0 + \mu H,$$

the separation $\Delta E = 2\,\mu H$ corresponding to a frequency $\nu = \Delta E/h = 2\mu H/h$. Inserting the value for the electron's magnetic moment (1 Bohr magneton) we have numerically $\nu = 2{\cdot}8\,H$ Mc/s with H expressed in oersteds. (Thus, if $H = 10\,000$ oersteds this corresponds to a wavelength of about 1 cm.)

Now suppose we were to reverse the magnetic field so quickly that the levels (1) and (2) are interchanged before the populations have time to change—we should then have

achieved the desired result. This turns out, however, to be very difficult in practice and as we change the magnetic field the populations follow. There *is*, however, a practical way in which the same result may be achieved. If we were to keep the field steady and apply a strong microwave signal at the resonant frequency, we should saturate the levels. If, however, the frequency is first of all well below resonance and is swept up through resonance to a value well above resonance it will be found that the populations of the two levels have been inverted. This was shown by Bloch (1946) for nuclear spins in his original and now classical paper on nuclear induction; it is, however, a quite general result for any two well-separated levels. Instead of sweeping the frequency, the magnetic field may be raised slightly in value to take the system across the resonance. The rate of sweep must be fast enough to accomplish the switch over in a time short compared with the relaxation time τ_1, but must also be slow enough to make the change-over 'adiabatic'. This latter condition comes from the analysis and is equivalent to saying that the internal molecular readjustment must be able to follow the changes. The method is known as 'adiabatic fast passage' and has been used by Combrisson, Honig and Townes (1956) to obtain stimulated emission using doped silicon. In this case the relaxation time τ_1 is very long at low temperatures. Sufficient emission was obtained to reduce appreciably the Q of the cavity, but not sufficient to obtain oscillations. *Maser* action has now been observed in doped Si (Feher *et al.* 1958) and also in a number of other materials (Chester, Wagner and Castle 1958) using this method of inversion. A pulsed amplifier might be made using the scheme indicated in Fig. 5; in this case the magnetic field is varied. At time T_1 the magnetic field H is set at a

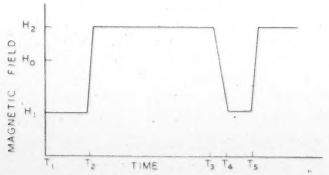


Fig. 5. Scheme for amplifier using adiabatic fast passage.

value H_1; a strong microwave signal is applied of frequency ν_0 given by $h\nu_0 = 2\mu H_0 (H_1 < H_0)$. At time T_2 the field is raised to the value $H_2 (H_2 > H_0)$ passing through the resonant value H_0. It is then held constant and the system should act as an amplifier at frequency $\nu = 2\mu H_2/h$. This it does till time T_3, provided $T_3 < \tau_1$, when the magnetic field is decreased, reaching the value H_1 at time T_4 when again it is held steady till time T_5, long enough to bring the spin system into thermal equilibrium. (For a substance like silicon, $T_5 - T_4$ may be quite short since optical injection of electrons may be used to decrease the relaxation time during this period.) After time T_5 the cycle is repeated. The amplifier is only active for a fraction of the time $(T_3 - T_2)/(T_5 - T_2)$, but this may be quite large.

There is also another, and in some ways more subtle, way of achieving inversion of energy levels. The transition probability theory which we have applied only really holds for weak stimulation. When strong microwave fields are involved we must use a direct quantum-mechanical treatment to derive the probability of a transition taking place from level (1) to level (2) or from level (2) to level (1). In this case the probability is no longer a linear function of the time. If the probability of excitation becomes large, clearly the molecule will almost certainly be excited to level (2) and then when in level (2) will be stimulated back to level (1), etc. We should therefore expect the probability of finding an atom in level (2) to be a periodic function of the time if we neglect relaxation, and this is just what is found. Suppose we know that at time $t = 0$ a molecule is in state (1). When stimulated at the resonant frequency it turns out that the probability $P(t)$ of finding it in state (2) is given by

$$P(t) = \sin^2\left(\frac{\pi W_i t}{h}\right) \tag{34}$$

where W_i is the interaction energy with the *peak* applied field. For an electric dipole $W_i = \rho \mathscr{E}$, where ρ is the dipole moment and $\mathscr{E}$ is the amplitude of the applied field, and for a magnetic dipole $W_i = \mu H$. For a waveguide carrying power P per unit area we have approximately

$$P = H^2 c/8\pi.$$

If $P = 1$ watt this gives $H = \frac{1}{10}$ oersted, so that $W_i/h = 0{\cdot}28$ Mc/s. The time T required to invert the levels is therefore given by

$$T = h/2W = 1{\cdot}8\ \mu\text{s}.$$

By means of a carefully controlled pulse of microwave power the levels could therefore be inverted. We have so far neglected relaxation effects, but if $\tau_1 \gg h/2W_i$ then clearly they will be unimportant. This method, so far as is known, has not yet been made to work in a *Maser*.

12. The three-level *Maser*

The most successful method of obtaining stimulated emission using a solid is that proposed by Bloembergen (1956). This employs three levels (1), (2), (3), having energies E_1, E_2, E_3 in ascending order (see Fig. 6). Now suppose we apply a strong microwave signal of frequency $\nu_{13} = (E_3 - E_1)/h$ so as to saturate the levels (1) and (3). We suppose the relaxation time τ_{13} for these two levels to be sufficiently long for this to be possible, so that we have $N_1 \simeq N_3$. Next we

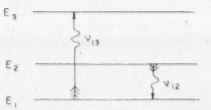


Fig. 6. Energy levels for 'three-level' *Maser*.

suppose that the levels (2) and (3) have a relaxation time τ_{23} so short that it is difficult to disturb the equilibrium ratio of their populations, without the presence of a strong signal of frequency $\nu_{23} = (E_3 - E_2)/h$. We then have

$$N_3/N_2 = \exp(-h\nu_{23}/kT). \tag{35}$$

Since $N_1 \simeq N_3$ we now have

$$N_1/N_2 = \exp(-h\nu_{23}/kT) \tag{36}$$

i.e. $N_1 < N_2$. If we had simply inverted the levels (1) and (2) we should have had

$$N_1/N_2 = \exp(-h\nu_{12}/kT). \tag{37}$$

We may write Eqn (36) in this form, namely

$$N_1/N_2 = \exp(-h\nu_{12}/kT_e) \tag{38}$$

where $T_e = T\nu_{12}/\nu_{23}$ and is called the effective inversion temperature. Thus we have when $h\nu_{12} \ll kT$

$$(N_2 - N_1)/N \simeq h\nu_{12}/2kT_e \tag{39}$$

this being the fraction of the molecules which effectively takes part in emission.

Maser amplifiers using this principle have now been used for a number of applications and provide receivers with effective noise temperatures of only a few degrees K (see § 15). The temperature T is frequently about 2° K, being provided by liquid helium pumped so as to be below its λ-point. The advantage of using as low a temperature as possible will be seen from Eqn (39) since the smaller the value of T the greater the fraction of 'active' molecules which can take part in stimulated emission. Ruby is the most commonly used material, in the form of single crystals of Al_2O_3 with about 0·1% of Cr. The Cr^{3+} ions in the crystal provide the active centres, being paramagnetic ions with spin quantum number 3/2. They would therefore provide *four* energy levels of which we could choose any three for *Maser* operation. (These would have energies $E_0 \pm \mu H$, $E_0 \pm 3\mu H$ if it were not for the effect of the crystalline field which modifies the variation of energy with magnetic field. The variation depends on the orientation of the crystal, and for ruby is now known in great detail.) A *Maser* amplifier using ruby has also been made to operate at a temperature obtained with liquid nitrogen, though naturally with increased effective noise temperature (Ditchfield and Forrester 1958). The pump power is normally provided at a frequency ν_{13} which is greater than the signal frequency ν_{12}. Recent developments, however, indicate that a subharmonic of ν_{13} might also be used, giving a pump frequency *lower* than the signal frequency. A number of other materials have been used particularly in attempts to obtain operation with signal frequency in the millimetre wave-band, but for most purposes ruby is still preferred.

In the first group of three-level *Masers* to operate, the active material was placed in a resonant cavity. In a more recent development the material is in the form of a strip in a travelling-wave tube. This arrangement has certain advantages, particularly as regards stability of operation (De Grasse, Schulz-DuBois and Scovil 1959). The great virtue of the 'solid-state' three-level *Maser* is that it has a fairly large bandwidth, of the order of a few Mc/s, and may be operated at low temperatures so as to obtain a very low effective noise temperature (see § 15).

A recent development of very great interest is the extension of this principle to provide amplification of light. In this case the frequency ν_{12} lies in the red part of the visible spectrum, the material used again being ruby. The energy levels are, however, some of the natural levels of the chromium ion and not those created by means of a magnetic field (Maiman 1960, Collins *et al.* 1960). Light amplifiers will be discussed in a later article.

14. Noise factors of *Masers*

We shall now give a very elementary discussion of the noise factors or effective noise temperatures obtainable with *Maser* amplifiers. In order to simplify the treatment we shall assume rather ideal conditions. The results of such a calculation, however, seem to be fairly general (Wittke 1957). We shall also neglect, to begin with, the losses in the waveguide; they may readily be taken into account, but complicate the algebra. Suppose we have noise power P_n and signal

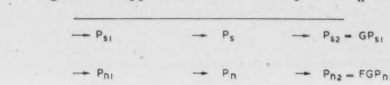


Fig. 7. Simplified waveguide *Maser*.

power P_s flowing down a waveguide (Fig. 7) filled with active material. Then we have

$$dP_s = J(N_2 - N_1)P_s dx \tag{40}$$

where J is a constant (equal to $h\nu a$ in Eqn (21)) and N_1 and N_2 have their usual meaning. Also we have

$$dP_n = J(N_2 - N_1)P_n dx + KN_2 dx. \tag{41}$$

The second term in (41) arises from what we have called spontaneous emission and is proportional to N_2, the number of molecules in the upper state. Now let the input and output signal and noise powers be respectively P_{s1}, P_{n1} and P_{s2}, P_{n2}. First, let us consider the case of thermal equilibrium when $P_s = 0$ and the source, waveguide and active material are all at the same temperature T_g. We shall denote this condition by zero suffixes. We have then $dP_n = 0$ so that

$$J(N_{20} - N_{10})P_{n0} + KN_{20} = 0.$$

Thus we have

$$K = JP_{n0}\left(\frac{N_{10}}{N_{20}} - 1\right) \simeq JP_{n0}h\nu/kT_g$$

but $P_{n0} = kT_g$ so we have

$$K = Jh\nu. \tag{42}$$

This determines the constant K in terms of J. We have then, replacing K in Eqn (41),

$$\frac{dP_n}{dx} = JN_2(P_n + h\nu) - JN_1P_n. \tag{43}$$

Integrating Eqn (40) we have, if l is the length of the guide occupied by active material,

$$\frac{P_{s2}}{P_{s1}} = \exp\{J(N_2 - N_1)l\} = G \tag{44}$$

where G is the power gain. For simplicity we shall suppose $G \gg 1$. Eqn (43) may also be readily integrated to give

$$P_{n2} = GP_{n1} + h\nu\left(\frac{N_2}{N_2 - N_1}\right)(G - 1). \tag{45}$$

When $G \gg 1$ the noise factor F is then given by

$$F = 1 + \frac{h\nu}{P_{n1}}\frac{N_2}{N_2 - N_1} \tag{46}$$

$$= 1 + \frac{h\nu}{kT_s}\frac{N_2}{N_2 - N_1} \tag{47}$$

where T_s is the temperature of the load. (A more general expression, applicable when G is not very much greater than 1 and taking into account guide losses, is given by Wittke (1957).) Eqn (47) gives the noise factor available in an ideal *Maser* and if the gain is high it may be closely approximated. The effective noise temperature T_r is obtained from Eqn (4) and is given by

$$T_r = N_2h\nu/k(N_2 - N_1). \tag{48}$$

Let us now examine various particular cases.

(i) *Separation Maser.*

In this case $N_1 = 0$ so we have

$$F = 1 + h\nu/kT_s \tag{49}$$

$$T_r = h\nu/k \tag{50}$$

with $\nu = 23\,870$ Mc/s and $T_s = 300°$ K, $F = 1{\cdot}004$ and $T_r = 1{\cdot}4°$ K. Values close to these have been observed, but it is clearly very difficult to give an exact figure.

(ii) *Inversion Maser.*

In this case $$\frac{N_2 - N_1}{N_2} = \frac{h\nu}{kT_g}$$

where T_g is the temperature of the waveguide or cavity, so we have

$$F = 1 + T_g/T_s \tag{51}$$

$$T_r = T_g. \tag{52}$$

If $T_g = T_s$, $F = 2$ (3 dB), while if $T_g = 3°$ K and $T_s = 300°$ K, $F = 1{\cdot}01$ and $T_r = 3°$ K.

(iii) *Three-level Maser.*

In this case we have (see Eqn (38)),

$$\frac{N_2 - N_1}{N_2} = \frac{h\nu}{kT_e} = \frac{h\nu}{kT_g}\frac{\nu_s}{\nu_p - \nu_s} \tag{53}$$

where ν_p and ν_s are the pump and signal frequencies. Hence we have

$$F = 1 + \frac{\nu_s}{\nu_p - \nu_s}\frac{T_g}{T_s} \tag{54}$$

$$T_r = \nu_s T_g/(\nu_p - \nu_s). \tag{55}$$

If $\nu_s \ll \nu_p$ we have $F \simeq 1$ even when $T_g = T_s$. For $\nu_s = \frac{1}{2}\nu_p$ we have

$$F = 1 + T_g/T_s \tag{56}$$

$$T_r = T_g \tag{57}$$

as for the inversion *Maser*.

For the conventional definition of the noise factor we have seen that T_s is normally taken as 300° K. In order to compare the performance of a number of receivers when used with an input corresponding to a given value of T_s we may either use the general equation for F or the effective noise temperature T_r. If for the *given* value of T_s the value of F obtained from Eqn (48) is appreciably greater than unity the receiver is not as good as it could be. Another way of saying this is that T_r should be much less than T_s if the receiver itself is not to contribute an appreciable amount of noise. On the other hand, there is little point in seeking to reduce the value of T_r if T_s is already much higher; for example, there is no point in using a receiver for which T_r is a few degrees K with an aerial and lossy waveguide system for which the effective aerial temperature is of the order of 100° K. With such a receiver it would be well worth while looking at really cold sources. This is perhaps the main application envisaged for the *Maser* amplifier, the most usual cold source being outer space. We must stress that in the above treatment no account has been taken of waveguide attenuation. If a source of temperature T_s is connected to an amplifier by a length l of waveguide (passive) at temperature T_g and having attenuation α, the effective temperature T_s' of the source as seen by the amplifier will be given by

$$T_s' = T_s e^{-\alpha l} + T_g(1 - e^{-\alpha l}). \tag{58}$$

If $T_g \gg T_s$, αl must be kept quite small, otherwise $T_s' \gg T_s$. Care will therefore have to be taken over lossy waveguide runs and it may even be necessary to cool the waveguide in certain circumstances.

15. Some applications of *Maser* amplifiers

As we have seen above, the main advantages of a *Maser* amplifier will only be realized when the effective noise temperature of the input is low. At microwave frequencies the effective temperature of outer space is very low, but it increases rapidly as the wavelength increases; this is illustrated in Fig. 8, which shows the variation of effective noise temperature for an ideal aerial with no side-lobes pointed well above the horizon and neglecting atmospheric absorption.

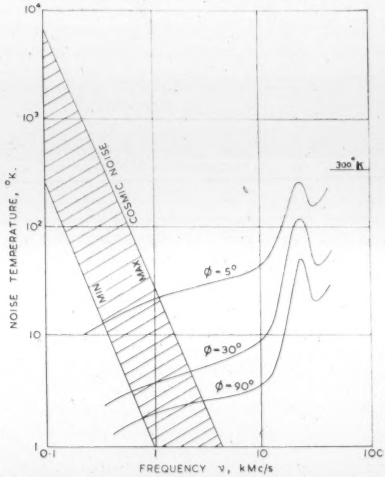


Fig. 8. Effective noise temperature of outer space and of the atmosphere as a function of wavelength (*after* Hogg 1959). ϕ = angle above horizontal.

A shaded band is shown to allow for the fact that the noise varies over the sky and the edges of the band correspond to maximum and minimum noise temperature. When account is taken of atmospheric absorption, the effective noise temperature of the atmosphere (also shown in Fig. 8) must be included (as with a lossy waveguide) and varies with the angle ϕ which the main lobe of the antenna makes with the horizon. If the antenna points nearly horizontally so that part of its main beam strikes the earth, the effective noise temperature will approach the earth temperature. If there are appreciable side lobes or backward 'spill-over' from the antenna, these will raise the effective noise temperature when the contribution from the atmosphere and outer space is less than 300° K. It will be seen that for frequencies above 0·3 Gc/s ($\lambda = 1$ m) the input noise is mainly from outer space and effective noise temperatures in excess of 300° K are encountered. At a wavelength of 3 cm, however, the main contribution comes from the atmosphere which for an aerial pointing at the zenith has an effective noise temperature of just over 3° K. For such wavelengths the use of a *Maser* amplifier is well worth while and has been used by Alsop

et al. (1958) for radio-astronomical observations, the *Maser* amplifier being mounted at the focus of the radio telescope. A *Maser* amplifier operating at a wavelength of 10 cm has also been used in conjunction with the radio telescope at the Royal Radar Establishment, Malvern. A *Maser* amplifier operating at frequencies near that of the hydrogen line at 1420 Mc/s has been used along with the Harvard radio telescope by Artman, Bloembergen and Shapiro (1958) to study galactic radiation from hydrogen. By far the most detailed analysis of the various contributions to the noise carried out so far for a low-noise system is that of De Grasse, Hogg *et al.* (1959 a, b), who combined a travelling-wave *Maser* amplifier with a specially designed antenna to obtain the remarkably low overall noise temperature of 18·5° K at a frequency of 5·65 Gc/s. The noise was made up as follows: the *Maser* amplifier had a gain of 35 dB and effective noise temperature of 10·5° K, the noise from outer space gave 2·5° K, aerial side lobes 2° K, antenna, waveguide and coupler loss 3·5° K. Low-noise systems of this type have recently been used at the Bell Telephone Laboratories in connection with the communications system depending on reflections from a satellite in the form of a spherical balloon.

16. Other forms of low-noise amplifiers

The *Maser* is not the only type of low-noise amplifier recently developed. Encouraged by the success of this type of amplifier designers of travelling-wave tubes have devised various means of reducing the noise which they generate. A type of amplifier known for a long time and depending for its operation on the cyclic variation of a reactive circuit element—the so-called parametric amplifier—has found practical application at very high frequencies through the use of variable reactance diodes available through semiconductor technology. Except when the very lowest possible noise temperature is desired such amplifiers may compete with the *Maser* type, especially at frequencies below 3000 Gc/s. They will be described in another article of this series.

17. Bibliography

A very large number of papers has appeared on the subject of *Masers* in the past five years and no attempt has been made to compile a complete list of references, only a small number of outstanding papers being quoted. Literature in book form is still rather scarce, but an introduction to the subject has been given by G. Troup (1959) and an extensive account of recent research in the form of a conference report has been published under the editorship of C. H. Townes (1960); this contains a considerable number of references to original papers.

References

Alsop, L. E., Giordmaine, J. A., Mayer, C. H., and Townes, C. H., 1958, *Astronom. J.*, **63**, 301.

Artman, J. O., Bloembergen, N., and Shapiro, S., 1958, *Phys. Rev.*, **109**, 1392.

Basov, N. G., and Prokhov, A. M., 1956, *Zh. Exper. Theor. Fiz.*, **30**, 3.

Bloch, F., 1946, *Phys. Rev.*, **70**, 460.

Bloembergen, N., 1956, *Phys. Rev.*, **104**, 324.

Chester, P. F., Wagner, P. E., and Castle, J. G., 1958, *Phys. Rev.*, **110**, 281.

Collins, R. J., Nelson, D. F., Schawlow, A. L., Bond, W., Garrett, C. G. B., and Kaiser, W., 1960, *Phys. Rev. Letters*, **5**, 303.

Combrisson, J., Honig, A., and Townes, C. H., 1956, *C.R. Acad. Sci., Paris*, **242**, 2451.

De Grasse, R. W., Hogg, D. C., Ohm, E. A., and Scovil, H. E. D., 1959a, *Proc. Nat. Electronics Conf.*, **15**, 370.

De Grasse, R. W., Schulz-DuBois, E. O., and Scovil, H. E. D., 1959b, *Bell Syst. Tech. J.*, **38**, 305.

Ditchfield, C. R., and Forrester, P. A., 1958, *Phys. Rev. Letters*, **1**, 448.

Feher, G., Gordon, J. P., Buehler, E., Gove, A. E., and Thurmond, C. D., 1958, *Phys. Rev.*, **109**, 221.

Goldenberg, H. M., Kleppner, D., and Ramsey, N. F., 1960, *Phys. Rev. Letters*, **5**, 361.

Gordon, J. P., Zeiger, H. J., and Townes, C. H., 1955, *Phys. Rev.*, **99**, 1264.

Hogg, D. C., 1959, *J. Appl. Phys.*, **30**, 1417.

Maiman, T. H., 1960, *Nature, Lond.*, **187**, 493.

Smith, R. A., Jones, F. E., and Chasmar, R. P., 1957, *The detection and measurement of infra-red radiation* (Oxford: University Press).

Townes, C. H. (Ed.), 1960, *Quantum Electronics* (Columbia: University Press).

Troup, G., 1959, *Masers* (London: Methuen Monographs).

Wittke, J. P., 1957, *Proc. Inst. Rad. Engrs, N.Y.*, **45**, 291.

The physics of computer elements

by C. N. W. LITTING, F.Inst.P., Electrical Engineering Laboratories, University of Manchester, Manchester 13

MS. received 13th February 1961

Abstract

The main requirements of a computing machine are that it should be able to receive and store information, perform logical operations on the information and finally to produce some form of useful output. Some interesting techniques employed in satisfying these requirements are described together with the physical phenomena on which they depend. The main use of novel techniques is in the storage of information and various systems are examined which depend on different physical principles. The following types of system are considered: acoustic, electrostatic, magnetic, optical and superconducting. It is considered that at present magnetic or superconducting devices employing thin films show most promise for the development of high speed devices.

1. Introduction

It is now nearly ten years since the first commercially available electronic digital computer was installed in this country (Williams and Kilburn 1951), and since that time many components employing interesting physical principles have been employed or proposed for use in computing machines.

In a digital computer, information, be it routing instructions or numbers, is fed along wires in the form of pulses of current or voltage. The main requirements of a machine could be enumerated as follows:

(i) Input of information—both numbers and instructions.
(ii) Storage of information.
(iii) Logical circuits to perform arithmetical operations.
(iv) Output of information.

The input and output operations are often performed by fairly conventional teleprinter tape, or punched card, equipment employing mechanical or photoelectric readers. Other interesting possible methods will be considered later.

The circuits to perform arithmetical operations are normally built up from a few simple logical 'gate' circuits in which thermionic valves or solid state devices are used as controlled switches.

Many methods have been employed for the storage of information, and make use of a large variety of physical principles. For the storage of binary digits any physical system which has two recognizable stable states may be used. In some cases it will be found that use may be made of the two stable states to produce more sophisticated input or output devices, or even to alter the logical arrangement of the machine.

It is therefore proposed to examine various types of storage depending on different physical properties and at the same time to mention any further use that can be made of these properties in a machine. Most of the devices considered can be classified into one of the following five types:

(i) Acoustic.
(ii) Electrostatic.
(iii) Magnetic.
(iv) Optical.
(v) Superconducting.

They will therefore be treated under these headings.

2. General requirements of a storage system

As has already been mentioned, any physical system which has two recognizable stable states could be used as the basis of a storage system. It is, of course, necessary to recognize or determine the state of the system in order to perform the operation of 'reading' the state of the system. It does not matter if the reading of the system destroys the stored information provided it can be re-written into the store immediately. To be of use in a computer, however, a store must have other properties as well. It must be possible to construct a large store containing many storage elements whilst still having rapid access for the reading or writing of information, at selected points in the store, without affecting the stored information at other points. Thus, storage systems often employ a two- or three-dimensional array of storage sites, a particular site being obtained by selection of the necessary coordinates.

3. Acoustic systems

One of the earliest types of storage system employed in computers was the mercury delay line (Wilkes and Renwick 1948). In this system the information to be stored is transmitted down a mercury tube in the form of acoustic waves: it is received at the far end at a later time and after reshaping of the pulses it is transmitted again. The system can, therefore, contain a certain amount of information in continuous circulation, the quantity being determined by the velocity of sound and the length of the path in the mercury.

As with all delay line systems of storage, this system suffers from the disadvantage that once the information has been transmitted down the line it cannot be made available until it has reached the other end, which for a five-foot length of mercury is approximately one millisecond later.

The original acoustic lines employed mercury and piezoelectric transducers, but use could be made of any suitable combination of transducer and material for the delay line. One example of another type is the use of a nickel wire delay line and magnetostrictive transducers (Millership *et al.* 1951).

4. Electrostatic systems

A storage system could be built employing individual capacitors to store the charge appropriate to every digit, but

this would be very expensive and bulky besides suffering from difficulties of access and digit selection. These difficulties were overcome in cathode-ray tube storage systems in which charge was stored on small areas of the insulating screen surface (Litting 1954).

The simplest of these systems, which was used in the early Manchester machines, employed a conventional cathode-ray tube (Williams and Kilburn 1949). In this system the information is stored as a charge pattern on the screen surface, each element of area being capacitively coupled to an external pick-up plate attached to the face of the tube. Thus, changing the potential of an element of area of the screen surface will result in an output from a high gain amplifier attached to the pick-up plate. To provide a storage system it is therefore necessary to set up the required charge pattern on the screen surface and to determine the pattern at a later time. This is accomplished by means of secondary emission produced by the cathode-ray tube beam.

The energy of the electrons in the beam is such that the secondary emission ratio of the screen surface is greater than unity. In these circumstances the whole of the screen surface acquires an equilibrium potential such that the net number of secondary electrons leaving the screen is equal to the number in the primary beam, the remainder of the secondary electrons returning to the surface. This potential is approximately equal to the potential of the internal conducting coating on the tube. At any time the actual spot bombarded by the beam will be a few volts positive with respect to the rest of the screen due to the velocity of emission of the secondaries, and the screen surrounding the spot will charge slightly negative due to the rain of returning secondaries. A positive 'well' is thus excavated on the surface, the depth of the well being determined by the velocity distribution of the secondaries (Fig. 1(*a*)).

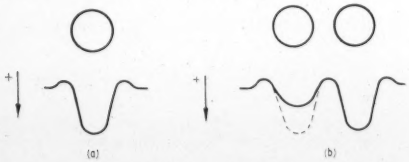


Fig. 1. Potential distribution on cathode-ray tube screen: (*a*) single spot; (*b*) double spot.

If the beam is switched on and off continuously there is no change in the charge on the screen surface, but every time the beam is switched on a cloud of electrons is produced in the vicinity of the screen. This cloud produces a transient pulse or cloud pulse at the output of the amplifier connected to the pick-up plate, and when the beam is switched off the cloud of electrons is removed and an opposite pulse is obtained. The output obtained when a single spot is switched on and off is shown in Fig. 2 (*a*) and (*b*).

If during the time when the beam is switched off a suitable square wave (Fig. 2(*c*)) is applied to the deflector plates, two separate spots will result.

If the separation of these two spots is one spot diameter, interaction takes place between the two resulting potential 'wells', the secondary electrons from one spot being attracted to the other spot. As these redistribution secondary electrons are of low kinetic energy they cause negligible secondary emission and hence tend to neutralize the original positive charge, i.e. they tend to 'fill' the well. Thus, when the beam impinges on the other spot, the well has to be re-excavated to re-attain the equilibrium potential. The effect of these processes which involve changes in the charge pattern on the screen surface are summarized in Fig. 1(*b*) and Fig. 2. The potential distribution indicated in Fig. 1(*b*) represents that pertaining to the instant at which the right-hand spot has been switched off. When the left-hand spot is now switched on, the left-hand well is excavated and, at a slower rate, the right-hand one is filled up. There are therefore three effects which contribute to the output of the amplifier: the cloud pulse, Fig. 2(*b*), a positive digging pulse, Fig. 2(*d*), representing the removal of negative charge, and a negative filling pulse, Fig. 2(*e*). The resultant output is thus of the form shown in Fig. 2(*f*). It is thus possible to leave a spot on the screen

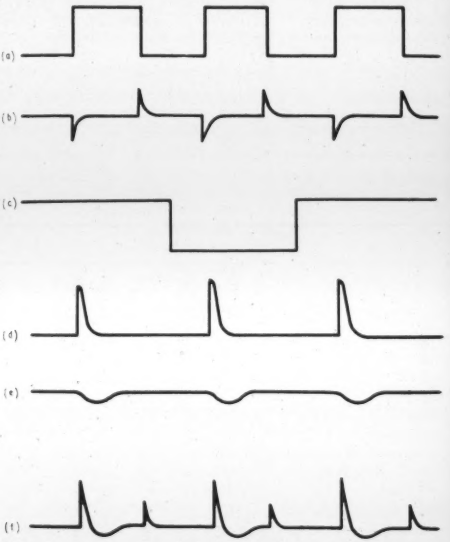


Fig. 2. Single and double-spot waveforms: (*a*) grid modulating waveform; (*b*) cloud pulses; (*c*) shift waveform; (*d*) digging pulses; (*e*) filling pulses; (*f*) amplifier output with double spot.

surface in a discharged or charged condition by merely bombarding the required spot and then bombarding or not bombarding an adjacent spot. The state of the spot can then be determined on subsequent bombardment, for a positive digging pulse will result if it had been discharged by an adjacent spot and a negative cloud pulse if it had not. We thus have the basis of a storage system, as the spot can be left in either of two states and identity of the state can be determined at a later time. In addition, the output may be used to operate a 'gate' circuit and 're-write' the desired charge distribution on the screen surface so that the stored information is not lost by 'reading' it. The stored charges gradually leak away, but of course they may be regenerated by merely reading and re-writing at appropriate intervals.

The system as described is one of the simplest forms of cathode-ray tube storage. More complicated storage tubes have been developed (Knoll and Kazan 1952). Some employ internal meshes which may be used to alter the equilibrium potential of the surface or to act as electrodes from which the

required signals may be picked up. By modulating the internal mesh electrode whilst 'writing' it is possible to leave the surface of the screen at two vastly different potentials. In some tubes these potentials are maintained indefinitely by means of a 'hold' beam and certain systems even employ a third beam for 'reading'.

It is also possible to use electron induced conductivity rather than secondary emission as the basis of a storage system, just as in the Vidicon television camera photoconductivity is used whilst the Emitron employs photoemission.

All forms of cathode-ray tube storage are slow by modern standards and are not normally employed in modern machines. For example the simple storage system described could not be operated at repetition frequencies greater than 250 kc/s.

It is perhaps worth noting that electrostatic charge storage combined with photoconductivity may be used to provide a xerographic output unit from a computer. If the output is displayed in a visible form (e.g. on a cathode-ray tube or electroluminescent panel) either in the form of a graph or a coded system of characters it may be permanently recorded by standard xerographic techniques (Huntley and Hughes 1959).

5. Magnetic systems

5.1. *Recording on moving magnetic material*

The magnetic properties of materials form the basis of many computing elements. The simplest use in a storage system is in the magnetic drum in which a cylindrical metal drum is coated with a suitable magnetic material on which information is stored in peripheral tracks in a somewhat similar way to a tape recording. A normal magnetic drum acts as a store for about 250 000 digits arranged in hundreds of parallel tracks. It thus forms a large capacity store but suffers from the disadvantage of having a large access time as it may take several milliseconds to read out the required information because it is necessary to wait for the drum to revolve till the required information reaches the output head.

Magnetic tape may also be used for high speed input and output from a machine but suffers from the disadvantage that it is not always easy to check that the required information has been written on the tape. To overcome this difficulty a 'static read head' has been developed for obtaining information off a stationary or slow-moving tape (Kilburn 1957). This head has an output sensitivity independent of tape speed and thus, by combining it with a conventional 'write' head, it is possible to check information character by character as it is written on the magnetic tape over a very wide range of tape speeds.

With conventional tape equipment the rate of change of magnetic flux in the 'read' head is proportional to the tape velocity and at low speeds the voltage output is therefore small or even zero. However, if the reluctance of the read head is modulated this will produce a rate of change of flux in the head even if the magnetized tape is stationary. By modulating the reluctance at a high frequency the rate of change of flux in the head, and hence the output voltage, is independent of the tape speed over a wide range which includes zero velocity. The output will be of twice the modulating frequency, its amplitude and phase being controlled by the stored signal on the tape.

The design of a static read head is shown in Fig. 3. It consists of a strip of Permalloy C bent into the shape of a head. At a point A opposite the gap B there is a hole through the centre of the material and a coil C is wound on the small toroid formed by this hole. The winding is such

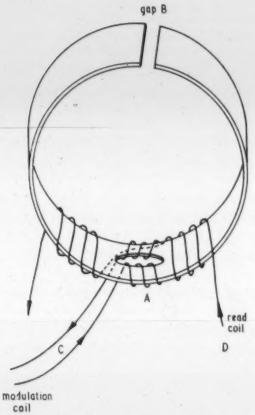


Fig. 3. The static read head.

that the flux from the coil passes round the small toroid but is balanced with respect to the normal magnetic path round the head. This coil is used to saturate the material surrounding the hole and so vary the reluctance of the head. On the other hand there is very little flux linkage between it and the read winding D. In addition, as the coil C produces no flux in the main part of the head, it cannot affect the information

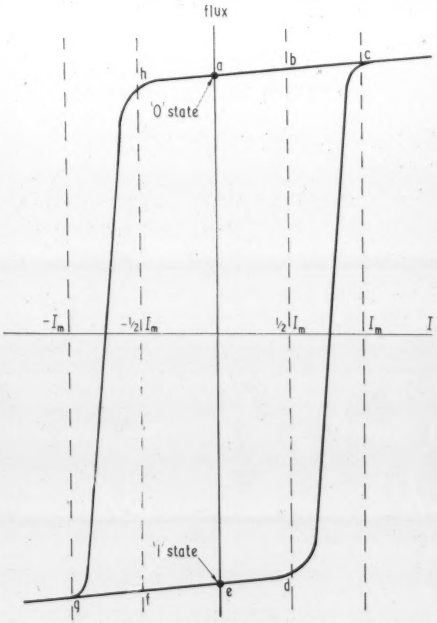


Fig. 4. Hysteresis loop of a magnetic core, plotted as total flux against current I linking it.

stored on the magnetic tape. Thus the only output from the read winding is that corresponding to the magnetic record on the tape.

5.2. *Magnetic cores*

In systems of storage employing magnetic cores the information is represented by the state of magnetization of the cores. One of the main difficulties in this system is the selection of the required information from a large store, and this has been overcome by making use of the properties of materials exhibiting so called 'square loop' hysteresis curves.

The approximately rectangular hysteresis loop of these materials enables the cores to be assembled in a three-dimensional array and to employ 'matrix' methods of selection. A typical hysteresis loop of one of these materials is shown in Fig. 4. In this loop the flux in the core is plotted against the current linking the core as this is of course proportional to the magnetizing force.

A current greater than or equal to I_m linking the core will saturate it in one direction and after that a current of $\frac{1}{2} I_m$ in the opposite direction will have negligible permanent effect on the magnetization. It is because of this property that it is possible to use matrix selection with core storage.

In a typical two-dimensional array (Robinson *et al.* 1956) the cores are threaded on wires as shown in Fig. 5. Any core can be magnetized in either direction without serious effect on

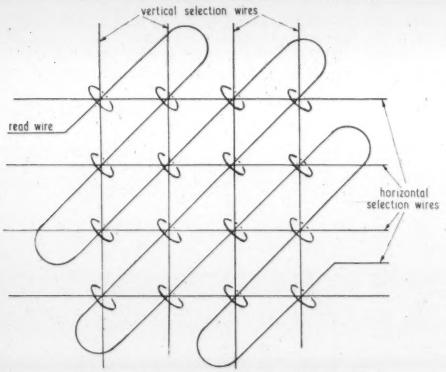


Fig. 5. Arrangement of cores in a matrix.

the other cores by passing a current of $\frac{1}{2} I_m$ in the correct direction through the horizontal and vertical selection wires linking it. The selected core then has the required current I_m linking it, but no other core suffers a permanent effect as the maximum current linking any other core is $\frac{1}{2} I_m$. By this means information may therefore be written into the store.

In order to read the information in a given core, currents of $\frac{1}{2} I_m$ are again passed through the appropriate selection wires to magnetize it to the '0' state (Fig. 4). This causes a pulse of voltage to be induced in the 'read' wire threading the core. The time integral of this pulse is proportional to the change in flux in the core. If the core is already in the '0' state there will be very little flux change (a to c in Fig. 4), whilst if it is in the '1' state there will be a much larger change (e to c). Thus from the presence or absence of a large output pulse on the read wire it is possible to determine whether a core is in the '1' or '0' state. It should be noted that in reading the state of a core by this method it is automatically left in the '0' state and therefore if it was originally in the '1' state steps will have to be taken to return it to the '1' state by the appropriate write pulses.

When reading a selected core all other cores in the same horizontal and vertical lines are linked by currents of $\frac{1}{2} I_m$ and these non-selected cores will thus induce small signals (flux changes a to b or e to d) in the read wires. However, it will be noted that the read wire is threaded in a diagonal manner through the array so that the signals from adjacent cores tend to cancel and so their sum is zero or negligible.

A finite time is required to switch a core from one state to another. This time is decreased if the driving current is increased or the size of core reduced. Using the type of selection system outlined above it is obviously impossible to increase the driving current and with the smallest cores currently available in this country it is difficult to design a system to give a complete read cycle time much less than 2 μs.

One possible solution of this difficulty is to use thin evaporated films of magnetic materials as storage elements. For example if a nickel–iron alloy (80% Ni 20% Fe) is evaporated on to a heated glass substrate in the presence of a magnetic field the films produced have rectangular hysteresis loops in the direction of the magnetic field. It seems probable that, using films of this type, cycle times of the order of 0·1 μs should be possible (Hoffman *et al.* 1960). At present one of the major difficulties is the reproducibility of the films as very little tolerance can be allowed with any matrix type of selection.

5.3. *A permanent store*

Another type of store which is required in a modern fast computer is one which contains a large amount of permanent or semi-permanent information and has a very rapid access time. This type of store may in fact be regarded as part of the logical system of the machine as it allows different interconnections of the logical circuits to be accomplished.

This requirement has been met by using the presence or absence of a rod of linear ferrite material to determine the state of the store (Kilburn and Grimsdale 1960). Consider the arrangement shown in Fig. 6, which consists of many

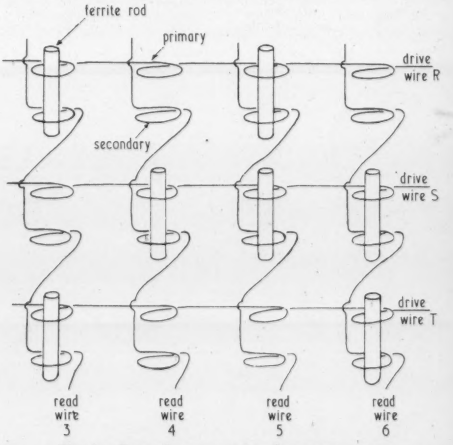


Fig. 6. General arrangement of permanent store.

transformers, with their primaries connected in series to form horizontal drive wires and their secondaries connected to form the vertical read wires. Each transformer represents a digit storage site, the presence of a ferrite rod denoting the '1' state and the absence of a rod the '0' state. When a drive pulse is passed along a horizontal drive wire an e.m.f. will be induced in the read wires which have ferrite rods at the appropriate intersections. For example, pulsing drive wire T in Fig. 6 will produce outputs in read wires 3 and 6 but not in wires 4 and 5.

The actual system that has been used has a capacity of 200 000 stored digits. It employs a woven mesh of tough enamel-covered wire, woven with wire of 33 s.w.g. and having approximately 18 meshes per inch. The drive wires are formed by adjacent pairs of the warp and the read wires by adjacent pairs of weft as shown in Fig. 7. In this figure the positions A, B, C and D are information sites and ferrite rods may be inserted in the mesh at these positions to represent stored digits. Such a rod is shown inserted at A. Similar rods are inserted in positions surrounding the storage sites. These rods or keepers (K in the figure) are to provide paths

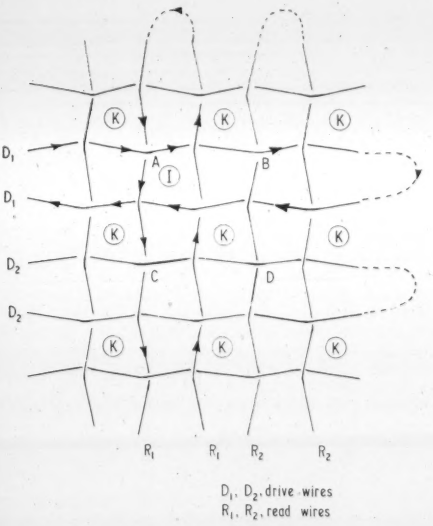


Fig. 7. Woven mesh store.

for the return flux from the information rods. This improves the magnetic circuit of the transformer giving an increased voltage output and also eliminates undesired coupling between adjacent information sites.

The complete store with 200 000 sites employs a sheet of woven mesh 54 in. by 8 ft which is folded back on itself to economize in space and for a store of this size a typical access time is 0·1 μs.

In an alternative version of the store there are 100 000 digit sites all of which can be changed in about 1 minute. This is achieved by mounting the ferrite rods in plastic cylinders and blowing them from one end of the tube to the other using a row of air jets which travel along the length of the store.

6. Optical systems

Two types of electro-optical devices are of interest in computer design, namely light sensitive and light producing devices. Normal photocells have long been used in optical tape readers, but in the last few years interest has grown in photoconductive and electroluminescent devices.

6.1. *Photoconductive devices*

The main drawback to the use of photoconductive devices of the cadmium sulphide type is their poor frequency response. For example it would be possible to build a fixed store consisting of horizontal drive wires and vertical read wires with a photoconductive connection at each intersection. The information would then be stored on punched cards in which a hole at the appropriate intersection would represent a '1' digit. Then on illuminating the device through the punched card and pulsing a given drive wire, outputs would only be obtained on the desired read wires.

The main difficulty of this system is the long access time due to the poor frequency response of the CdS. Another difficulty is the presence of cross-talk. This is due to the fact that unless the input impedance of the read wires are absolutely zero a voltage pulse will appear on them and because many junctions are illuminated they will produce voltages on other drive wires which will result in unwanted pulses on further read wires.

However, a new type of photodiode based on the combination of a photoconductor and an electret has been produced (Diemer and Van Santen 1960), and this removes the problem of cross-talk. The new element consists of a sprayed and sintered layer on which the necessary cross-bar conductors can be deposited and the photodiode connections formed at the interconnections by suitable processing. In the dark these interconnections then have a near infinite impedance but when illuminated have a diode characteristic. If now an a.c. signal is applied to a drive wire a rectified version appears on the desired read wires and cross-talk is eliminated as at least one of the junctions through which it must pass will be in the reverse direction.

6.2. *Electroluminescence*

The main use of electroluminescent devices would at present seem to be in the production of visual output devices or curve plotters. They can of course also be used in conjunction with photoconductive units to form two state devices although these would be slow in operation.

A typical electroluminescent cell consists of a layer of suitable phosphor powder embedded in transparent insulating binder and sandwiched between two conducting electrodes at least one of which is transparent. When an alternating voltage is applied between the two electrodes light output is obtained from the phosphor. The light output from the cell is a function of both the frequency and the amplitude of the applied voltage. At constant frequency the light output is governed by the expression

$$\text{Integrated light output} \quad B = a \exp(-b/V^{1/2})$$

where a and b are constants and V is the amplitude of the applied voltage.

A graphical output display panel consists of an electroluminescent layer with a series of horizontal electrodes on one side and vertical electrodes on the other (Kilburn 1957). This matrix or cross-bar system enables any spot on the panel to be brightened by applying a voltage between the appropriate pair of electrodes.

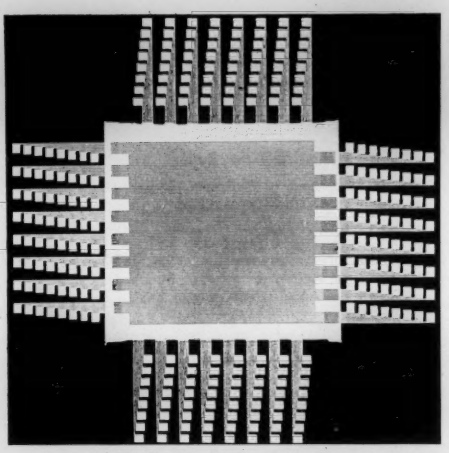


Fig. 8. Graphical output panel.

A typical small panel employing 128 by 128 conductors is shown in Fig. 8 (Kilburn *et al.* 1958).

One of the main advantages of this method of curve plotting compared with normal cathode-ray tube methods is its accuracy and reproducibility since the spot selection is still digital and the matrix has a fixed geometry.

In selecting a spot on the panel voltages of opposite sign are applied to the appropriate conductors. The selected intersection thus receives the desired voltage but all other points on the same horizontal and vertical rows receive half voltage and thus tend to emit light. However, it is easily shown from the expression governing the brightness that the ratio of brightness of the selected spot to the others is $\exp\{b(\sqrt{2}-1)/V\}$ where V is the voltage at the selected intersection. With an improved method of biasing the matrix it is possible to obtain a voltage discrimination ratio of three and hence a brightness discrimination ratio of $\exp\{b(\sqrt{3}-1)/V\}$ between the selected spot and the background.

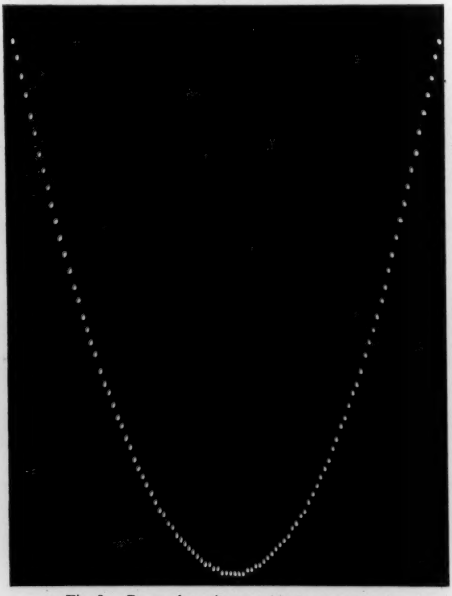


Fig. 9. Curve plot using graphical output panel.

Using typical values, which would enable a 1000 point graph to be plotted and photographed in 20 seconds, this ratio is 80000 and therefore the background illumination is negligible. A photograph of a typical curve plot is shown in Fig. 9 (Jeffreys 1960, unpublished).

An electroluminescent panel can also be used as part of an input system or permanent store in that it can be used for the scanning of punched cards. In such a system it would be used in conjunction with a single photosensitive device such as a photomultiplier and the selection of the required information would be performed with the panel rather than by using individual light detectors.

7. Superconducting systems

In 1911 Kamerlingh-Onnes discovered that if mercury is cooled in a magnetic field the temperature for transition to the superconducting state depends on the magnitude of the magnetic field, being lower the greater the magnetic field. This behaviour is similar to that for all superconducting metals and thus a strip of material may be switched from superconducting to normal by applying a sufficient magnetic field. This may be achieved by applying an external field, by means of a current in a coil or in a wire running parallel to the strip or by passing a current through the strip itself.

This phenomenon was first demonstrated as feasible for use in a computer element by the development of a switching device called the cryotron (Buck 1956). This device consists of a fine wire passing through the centre of a coil, but insulated from it. Both the coil and wire are of superconducting materials and the current in the coil is used to switch the wire to the resistive state.

Provided the device has a gain greater than unity, that is, the current the wire can carry without losing its superconductivity is greater than that required in the coil to return the wire to its resistive state, one unit can be used to provide the control current for other units.

Several interesting circuits were developed using these types of cryotron but they suffer from the basic disadvantage that the speed of operation of the wire-wound cryotron is low, being of the order of a few kilocycles per second, owing to the inductive time constant of the device.

Cryotrons have since been produced by vacuum evaporation techniques with switching times less than $0{\cdot}1\ \mu s$ (Smallman *et al.* 1960).

In these thin film cryotrons the wire to be switched is replaced by a thin film of tin whilst the coil is replaced by an evaporated strip of lead across the tin. The lead, which has a higher critical temperature than the tin, is insulated from it by an evaporated insulating film, and a current in the lead is used to control the state of the tin film. This type of construction allows compact circuits to be produced: for example, the circuit of a five-stage flip–flop ring oscillator employing ten cryotrons was produced on a substrate measuring 1 in. by 2 in.

An alternative method of utilizing superconducting properties is to set up a persistent current in a loop in order to store a digit. Because of the zero resistance this current will continue to flow indefinitely and the direction of the flow can be used to indicate a '0' or '1' state. The current in the loop can be altered by means of a drive wire linked inductively with the superconducting circuit, but unless the current in the circuit is raised above the critical value, so that superconductivity is destroyed, the final current in the loop cannot be effected by a drive pulse which returns to its initial level. It is therefore possible to use a matrix type of selection system in which the required critical current is supplied by two drive wires coupled to the loop. A read wire also coupled to the loop will detect any change of current in, and therefore flux through, the loop.

In suggested elements of this type the storage units are produced by evaporation and the various write and read circuits are also insulated evaporated strips lying over the storage loop (Crowe 1957, Crittenden *et al.* 1960).

This type of construction would allow an extremely compact store with short access time to be produced. At present the difficulty is to obtain sufficiently reproducible evaporated films to enable a large matrix to be constructed. It seems probable that the small quantities of gas occluded during the evaporation of a film may affect its superconducting properties and this indicates the importance of the degree of vacuum obtained during the evaporation.

It is considered feasible that using these types of evaporated storage elements it should be possible to produce a store with a capacity of 10^7 digits with an access time of less than 50 μs in a cube of 40 cm sides.

There is of course the problem of keeping the elements cold, for whenever a storage loop becomes resistive power is dissipated. In a store of the type envisaged above, the repeated switching of a single storage loop at 20 Mc/s would result in a dissipation of the order of 2×10^{-14} watt, corresponding to the evaporation of 2 cm^3 of liquid helium per hour.

8. Conclusion

Various types of storage systems have been described depending for their action on different physical principles.

The systems embodying superconductive elements and evaporated magnetic films at present seem to offer most scope for development. With both these types of system it seems feasible that large storage units could be produced with access times much shorter than any other type of store.

However, when one considers the many and varied physical phenomena which have already been applied to the production of computer elements in just over a decade it would not be surprising if in the course of a few years new types of systems are developed depending for their action on other physical properties.

Acknowledgments

The author wishes to acknowledge his indebtedness to his colleagues in the Electrical Engineering Department and in particular to Professor T. Kilburn and the members of the Computing Machine Laboratory at Manchester University.

References

Buck, J. F., 1956, *Proc. Inst. Radio Engrs, N.Y.*, **44**, 482.

Crittenden, E. C., Jr., Cooper, J. N., and Schmidlen, F. W., 1960, *Proc. Inst. Radio Engrs, N.Y.*, **48**, 1233.

Crowe, J. W., 1957, *I.B.M. J. Res. Dev.*, **1**, 294.

Diemer, G., and Van Santen, J. G., 1960, *Philips Res. Rep.*, **15**, 368.

Hoffman, G. R., Turner, J. A., and Kilburn, T., 1960, *J. Brit. Instn Radio Engrs*, **20**, 31.

Huntley, K. G., and Hughes, J., 1959, *Proc. Instn Elect. Engrs*, Pt B, **106**, 454.

Kilburn, T., 1957, *Proceedings of Data Processing and Automatic Computing Machines Conference, Salisbury* (Adelaide Southern Australia, Weapons Research Establishment).

Kilburn, T., Hoffman, G. R., and Hayes, R. E., 1958, *Proc. Instn Elect. Engrs*, Pt B, **105**, 136.

Kilburn, T., and Grimsdale, R. L., 1960, *Proc. Instn Elect. Engrs*, Pt B, **107**, 567.

Knoll, M., and Kazan, B., 1952, *Storage Tubes* (New York: John Wiley).

Litting, C. N. W., 1954, *J. Sci. Instrum.*, **31**, 351.

Millership, R., Robbing, R. C., and De Barr, A. E., 1951, *Brit. J. Appl. Phys.*, **2**, 304.

Robinson, A. A., Newhouse, V. L., Friedman, M. J., Bindon, D. G., and Carter, I. P. V., 1956, *Proc. Instn Elect. Engrs*, Pt B, **103**, Suppl. 2, 295.

Smallman, C. R., Slade, A. E., and Cohen, M. L., 1960, *Proc. Inst. Radio Engrs, N.Y.*, **48**, 1562.

Wilkes, M. V., and Renwick, W., 1948, *Electronic Engng*, **20**, 208.

Williams, F. C., and Kilburn, T., 1949, *Proc. Instn Elect. Engrs*, Pt II, **96**, 183.

—— 1951, Manchester University Computer Inaugural Conference.

Emission of negative ions of oxygen from dispenser cathodes Part 1.—Cathodes of barium oxide in sintered nickel

by N. A. SURPLICE, B.Sc., Ph.D., A.Inst.P., Physics Department, University College of North Staffordshire, Keele, Staffs.

MS. received 14*th November* 1960

Abstract

Cathodes of sintered nickel and barium oxide have been used as ion sources in simple mass spectrometers and have been found to emit negative ions of atomic oxygen for at least 24 *hours during their activation at* 1250–1350° K. *The evidence suggests that the oxygen is produced by the dissociation of barium oxide and is removed from the cathode by positive ion bombardment. At voltages greater than the ionization potentials of the residual gases, the variation of the oxygen ion current I with time t could be expressed as the sum of two exponentially decreasing terms:* $I = a \exp(-pt) + b \exp(-qt)$, *where p and q were of the order of* 2 *and* 0·2 *per hour respectively. Similar results have been obtained with a commercial cathode containing triple carbonates. There appear to be two processes operating in parallel, corresponding to the two terms of the equation. It is postulated that the first process is the diffusion of oxygen through the bulk of the crystals to the nickel part of the cathode's surface, and its removal from there by bombardment with positive ions of oxygen and nitrogen. The second process is the diffusion of oxygen through the pores of the cathode, and its removal from the ends of the pores by bombardment with positive ions of barium as well as of the residual gases. The first process has the slower rate of replacement of oxygen and in an hour or two this process is insignificant compared with the second one. If the cathode's surface is temporarily poisoned with oxygen from an external source then it rids itself of the extra oxygen within a few seconds.*

1. Introduction

CATHODES made from a mixture of alkaline earth carbonates with powdered nickel are known as barium–nickel dispenser cathodes (Balas *et al.* 1955, Beck *et al.* 1954, Fane 1958, Macnair *et al.* 1953, Richardson 1957, Richardson and Vick 1960). The mixture is first pressed into a pellet, then heated *in vacuo* or in hydrogen to decompose the carbonates and to sinter the metal. At first the cathode does not give a sufficiently high emission current of electrons, but it can be made to do so by means of the activation treatment, which consists of drawing current from it for a few hours. The active cathode appears to consist of alkaline earth oxide grains within a metal matrix, whose work function has been reduced by at least a partial covering of barium (Fane 1958). The activation process presumably causes electron donors to form in the oxide lattice and barium atoms to diffuse over the metal matrix. The rates of diffusion and evaporation of barium and barium oxide have already been studied (Richardson and Vick 1960), and the rate of evolution of negative oxygen ions is described below.

2. Apparatus

The cathodes were made from only one alkaline earth oxide in order to keep the dispensing system as simple as possible, although in consequence they gave an average electron emission of only 50 mA/cm^2. They were made to Richardson's (1957) specification of 69·9% by weight of nickel powder of 4–5 μ diameter (containing 0·1% carbon, 0·02% iron, 0·001% sulphur), 30% barium carbonate, and 0·1% silicon. Samples of 0·45 g of the mixture were pressed into disks 0·1 cm thick which closed one end of O-nickel tubes of 0·5 cm internal diameter. A batch of ten cathodes was made at approximately the same pressure with a fly press, and a batch of sixteen was made by Dr. J. F. Richardson with a hydraulic press at Metropolitan-Vickers Electrical Co. Ltd. The cathodes were kept in a desiccator until they were used.

Each cathode in turn was made the ion source of one of three small mass spectrometers. Two of the instruments (types MS8 and MS9) could distinguish negative ions of masses up to 45 with a resolving power of about a hundred. The third (type MS10) had much wider entrance and exit slits and deflection chamber, so it had only half the range and resolving power of the other two, but double the beam current. The negative ions emitted from the cathode were accelerated by the modulator and anode of a simple electron gun, then selected by a magnetic analyser and detected by a d.c. amplifier at the collector (Fig. 1). Under typical operating conditions only 2×10^{-7} of the electron current from the

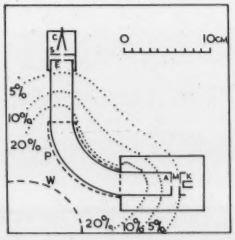


Fig. 1. Block diagram of mass spectrometer, showing fringing field of magnet. K, cathode; M, modulator; A, anode; E, exit slit; S, screen; C, collector; P, mild-steel pole piece; W, winding on yoke of magnet. Dotted lines show fringing field as a percentage of the field strength between the centres of the pole pieces.

cathode reached the collectors of MS8 and MS9, but it was estimated that perhaps 10^{-5} of the negative ion current would reach the collectors, because the ions would be deviated less than the electrons by the wide fringing field of the electromagnet. The mass spectrometers had glass envelopes, so that they could be outgassed by baking, and they have been described more fully elsewhere (Surplice 1960).

Each mass spectrometer was evacuated by a mercury diffusion pump, through a liquid air trap. It was first baked for an hour and its electrodes were eddy-current heated to redness. The barium carbonate was converted to oxide by gradually increasing the cathode temperature from 600–1200° K over three hours, and then the cathode was sintered for twenty minutes at 1300° K. The whole instrument was baked at 400° C for at least twelve hours. Finally, the cathode was kept at 1200° K while the other electrodes were eddy-current heated again, and then the pumping tube was sealed off and barium getters were fired. Throughout these operations the pressure was too low to be recorded by a Penning gauge so may be assumed to have been well below 10^{-5} mm Hg during breakdown and better than 10^{-6} mm Hg when the getters had been fired. The cathode temperature was measured with a tungsten–nickel thermocouple with its hot junction on the upper rim of the nickel tube that contained the cathode pellet. The cathodes of the first batch are referred to by the letters A–J, and those of the second batch by K–Z, so cathode 8A means cathode A in the mass spectrometer MS8.

3. Results

The mass spectrum

Each cathode was activated by drawing electron current from it, usually at 1250° K. The mass spectrum of the negative ions emitted by it during this process was studied at frequent intervals, and a typical example is shown in Fig. 2.

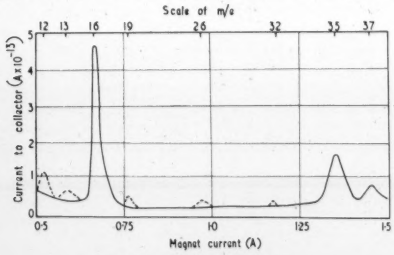


Fig. 2. The mass spectrum.
——— ions emitted by all cathodes.
– – – ions emitted by only a few cathodes.

Atomic oxygen and chlorine negative ions were present in the mass spectra of all cathodes. The peak currents to the collector when it was a few volts positive with respect to the cathode were of the order of 10^{-12} A of atomic oxygen and 10^{-13} A of chlorine. The mass spectra of several cathodes showed a small peak of about 3×10^{-14} A of an ion of mass number 26, but only one showed the molecular oxygen ion $^{32}O^-$ and only one other showed oxides of carbon $^{28}CO^-$ and $^{44}CO_2^-$. The heights of the ion current peaks were proportional to their areas and were therefore a satisfactory measure of the ion current.

In some of the experiments with tube MS8 the chlorine ion currents were increased thirty-fold by putting a spot of lithium chloride solution on to the electron gun, and the chlorine ions were then used to calibrate the mass spectrum. Another peak of about 5×10^{-14} A of $^{19}F^-$ appeared with the increased chlorine peaks and presumably came from a fluoride impurity in the lithium chloride, but no negative ions of lithium were seen. The source of the original chlorine ions has been shown to be a chloride flux used in the manufacture of the glass envelope of the mass spectrometer (Hamaker *et al.* 1947, Vick and Walley 1954). The ion of mass number 26 was probably acetylene, which had been produced by a carbide reacting with water vapour. The carbide might have come either from the 'Kemet' K.I.C. getters or from the reaction of barium with the carbonyl nickel powder in the cathode. Small peaks of the ions $^{12}C^-$ and $^{13}CH^-$ appeared in the mass spectrum after experiments on poisoning, owing to the method of poisoning the cathode (Surplice 1960).

A special cathode of nickel powder was used to make sure that the source of the oxygen was the barium oxide, rather than an adsorbed layer of oxygen on the nickel. This cathode was made in the same way as all the others, except for the omission of barium carbonate, and it was taken through precisely the same procedure of heating from 600–1200° K, baking the tube and sintering at 1300° K. It was also activated in the same way, but although for a few hours its electron emission was comparable with that of the others it showed no trace at all of oxygen negative ion emission, either at 1250° K or 1350° K. The electron emission was presumably an effect of the impurities in the nickel powder; it began as 6 mA at 1250° K and 250 V, but fell to 3 mA in 24 hours. This electron emission was adequate to ionize the residual gas around the cathode, and there would have been no lack of positive ions to strip any adsorbed layer of oxygen off the nickel.

The energies of the ions

The energy distribution of each type of ion was found by measuring its current to the collector against a series of

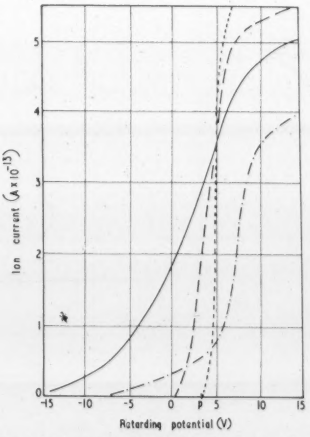


Fig. 3. The energy distribution for various ions.
——— $^{16}O^-$
– – – $^{35}Cl^-$
—·—· $^{19}F^-$
..... electrons

Emission of negative ions of oxygen from dispenser cathodes Part 1.—Cathodes of barium oxide in sintered nickel

by N. A. SURPLICE, B.Sc., Ph.D., A.Inst.P., Physics Department, University College of North Staffordshire, Keele, Staffs.

MS. received 14*th November* 1960

Abstract

Cathodes of sintered nickel and barium oxide have been used as ion sources in simple mass spectrometers and have been found to emit negative ions of atomic oxygen for at least 24 *hours during their activation at* 1250–1350° K. *The evidence suggests that the oxygen is produced by the dissociation of barium oxide and is removed from the cathode by positive ion bombardment. At voltages greater than the ionization potentials of the residual gases, the variation of the oxygen ion current I with time t could be expressed as the sum of two exponentially decreasing terms:* $I = a \exp(-pt) + b \exp(-qt)$, *where p and q were of the order of* 2 *and* 0·2 *per hour respectively. Similar results have been obtained with a commercial cathode containing triple carbonates. There appear to be two processes operating in parallel, corresponding to the two terms of the equation. It is postulated that the first process is the diffusion of oxygen through the bulk of the crystals to the nickel part of the cathode's surface, and its removal from there by bombardment with positive ions of oxygen and nitrogen. The second process is the diffusion of oxygen through the pores of the cathode, and its removal from the ends of the pores by bombardment with positive ions of barium as well as of the residual gases. The first process has the slower rate of replacement of oxygen and in an hour or two this process is insignificant compared with the second one. If the cathode's surface is temporarily poisoned with oxygen from an external source then it rids itself of the extra oxygen within a few seconds.*

1. Introduction

CATHODES made from a mixture of alkaline earth carbonates with powdered nickel are known as barium–nickel dispenser cathodes (Balas *et al.* 1955, Beck *et al.* 1954, Fane 1958, Macnair *et al.* 1953, Richardson 1957, Richardson and Vick 1960). The mixture is first pressed into a pellet, then heated *in vacuo* or in hydrogen to decompose the carbonates and to sinter the metal. At first the cathode does not give a sufficiently high emission current of electrons, but it can be made to do so by means of the activation treatment, which consists of drawing current from it for a few hours. The active cathode appears to consist of alkaline earth oxide grains within a metal matrix, whose work function has been reduced by at least a partial covering of barium (Fane 1958). The activation process presumably causes electron donors to form in the oxide lattice and barium atoms to diffuse over the metal matrix. The rates of diffusion and evaporation of barium and barium oxide have already been studied (Richardson and Vick 1960), and the rate of evolution of negative oxygen ions is described below.

2. Apparatus

The cathodes were made from only one alkaline earth oxide in order to keep the dispensing system as simple as possible, although in consequence they gave an average electron emission of only 50 mA/cm^2. They were made to Richardson's (1957) specification of 69·9% by weight of nickel powder of 4–5 μ diameter (containing 0·1% carbon, 0·02% iron, 0·001% sulphur), 30% barium carbonate, and 0·1% silicon. Samples of 0·45 g of the mixture were pressed into disks 0·1 cm thick which closed one end of O-nickel tubes of 0·5 cm internal diameter. A batch of ten cathodes was made at approximately the same pressure with a fly press, and a batch of sixteen was made by Dr. J. F. Richardson with a hydraulic press at Metropolitan-Vickers Electrical Co. Ltd. The cathodes were kept in a desiccator until they were used.

Each cathode in turn was made the ion source of one of three small mass spectrometers. Two of the instruments (types MS8 and MS9) could distinguish negative ions of masses up to 45 with a resolving power of about a hundred. The third (type MS10) had much wider entrance and exit slits and deflection chamber, so it had only half the range and resolving power of the other two, but double the beam current. The negative ions emitted from the cathode were accelerated by the modulator and anode of a simple electron gun, then selected by a magnetic analyser and detected by a d.c. amplifier at the collector (Fig. 1). Under typical operating conditions only 2×10^{-7} of the electron current from the

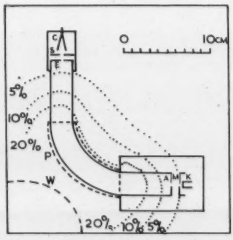


Fig. 1. Block diagram of mass spectrometer, showing fringing field of magnet. K, cathode; M, modulator; A, anode; E, exit slit; S, screen; C, collector; P, mild-steel pole piece; W, winding on yoke of magnet. Dotted lines show fringing field as a percentage of the field strength between the centres of the pole pieces.

cathode reached the collectors of MS8 and MS9, but it was estimated that perhaps 10^{-5} of the negative ion current would reach the collectors, because the ions would be deviated less than the electrons by the wide fringing field of the electromagnet. The mass spectrometers had glass envelopes, so that they could be outgassed by baking, and they have been described more fully elsewhere (Surplice 1960).

Each mass spectrometer was evacuated by a mercury diffusion pump, through a liquid air trap. It was first baked for an hour and its electrodes were eddy-current heated to redness. The barium carbonate was converted to oxide by gradually increasing the cathode temperature from 600–1200° K over three hours, and then the cathode was sintered for twenty minutes at 1300° K. The whole instrument was baked at 400° C for at least twelve hours. Finally, the cathode was kept at 1200° K while the other electrodes were eddy-current heated again, and then the pumping tube was sealed off and barium getters were fired. Throughout these operations the pressure was too low to be recorded by a Penning gauge so may be assumed to have been well below 10^{-5} mm Hg during breakdown and better than 10^{-6} mm Hg when the getters had been fired. The cathode temperature was measured with a tungsten–nickel thermocouple with its hot junction on the upper rim of the nickel tube that contained the cathode pellet. The cathodes of the first batch are referred to by the letters A–J, and those of the second batch by K–Z, so cathode 8A means cathode A in the mass spectrometer MS8.

3. Results

The mass spectrum

Each cathode was activated by drawing electron current from it, usually at 1250° K. The mass spectrum of the negative ions emitted by it during this process was studied at frequent intervals, and a typical example is shown in Fig. 2.

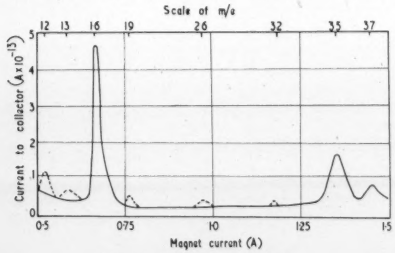


Fig. 2. The mass spectrum.
——— ions emitted by all cathodes.
– – – ions emitted by only a few cathodes.

Atomic oxygen and chlorine negative ions were present in the mass spectra of all cathodes. The peak currents to the collector when it was a few volts positive with respect to the cathode were of the order of 10^{-12} A of atomic oxygen and 10^{-13} A of chlorine. The mass spectra of several cathodes showed a small peak of about 3×10^{-14} A of an ion of mass number 26, but only one showed the molecular oxygen ion $^{32}O^-$ and only one other showed oxides of carbon $^{28}CO^-$ and $^{44}CO_2^-$. The heights of the ion current peaks were proportional to their areas and were therefore a satisfactory measure of the ion current.

In some of the experiments with tube MS8 the chlorine ion currents were increased thirty-fold by putting a spot of lithium chloride solution on to the electron gun, and the chlorine ions were then used to calibrate the mass spectrum. Another peak of about 5×10^{-14} A of $^{19}F^-$ appeared with the increased chlorine peaks and presumably came from a fluoride impurity in the lithium chloride, but no negative ions of lithium were seen. The source of the original chlorine ions has been shown to be a chloride flux used in the manufacture of the glass envelope of the mass spectrometer (Hamaker *et al.* 1947, Vick and Walley 1954). The ion of mass number 26 was probably acetylene, which had been produced by a carbide reacting with water vapour. The carbide might have come either from the 'Kemet' K.I.C. getters or from the reaction of barium with the carbonyl nickel powder in the cathode. Small peaks of the ions $^{12}C^-$ and $^{13}CH^-$ appeared in the mass spectrum after experiments on poisoning, owing to the method of poisoning the cathode (Surplice 1960).

A special cathode of nickel powder was used to make sure that the source of the oxygen was the barium oxide, rather than an adsorbed layer of oxygen on the nickel. This cathode was made in the same way as all the others, except for the omission of barium carbonate, and it was taken through precisely the same procedure of heating from 600–1200° K, baking the tube and sintering at 1300° K. It was also activated in the same way, but although for a few hours its electron emission was comparable with that of the others it showed no trace at all of oxygen negative ion emission, either at 1250° K or 1350° K. The electron emission was presumably an effect of the impurities in the nickel powder; it began as 6 mA at 1250° K and 250 V, but fell to 3 mA in 24 hours. This electron emission was adequate to ionize the residual gas around the cathode, and there would have been no lack of positive ions to strip any adsorbed layer of oxygen off the nickel.

The energies of the ions

The energy distribution of each type of ion was found by measuring its current to the collector against a series of

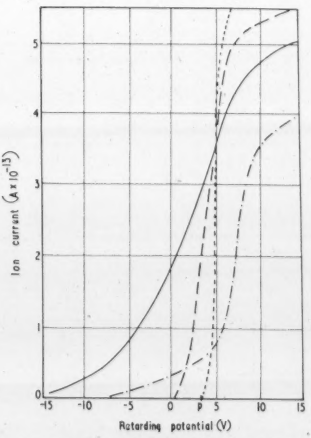


Fig. 3. The energy distribution for various ions.
——— $^{16}O^-$
– – – $^{35}Cl^-$
—·—· $^{19}F^-$
..... electrons

retarding potentials. As the retarding potential was increased to allow only ions of high energy and momentum to be collected, so the magnetic field was similarly increased to keep the ion beam focused on the exit slit. Some examples of the energy distributions for various ions are shown in Figs 3 and 4. In Figs 3 and 4 the point p, where the energy distribution of the electrons cuts the abscissa, represents the contact potential difference between the surface of the cathode and the nickel collector. The parts of the curves to the left of p represent ions that were given extra energy by being sputtered off the cathode by positive ion bombardment. The parts of the curves to the right of p represent ions that had been formed in the residual gas between the cathode and the rest of the electron gun. The point where the retarding potential was zero had no particular significance for these cathodes, because their surface was an intimate mixture of nickel and barium oxide.

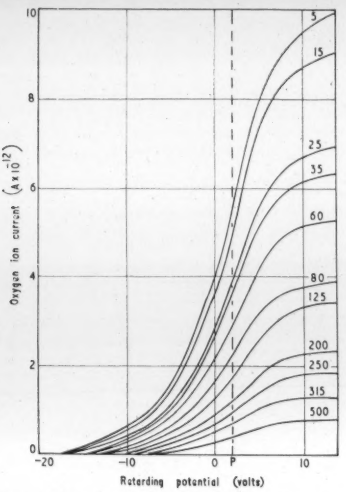


Fig. 4. Change of oxygen ion current with time (cathode 8B). Numbers on curves refer to time in minutes after the beginning of activation.

More than half the atomic oxygen ions were sputtered from each cathode: but only about 3% of the molecular oxygen ions were sputtered from the only cathode that emitted them at all. About 15% of the fluorine ions, but only a few of the chlorine ions, were sputtered from the cathode. The currents of the other ions were either too small or too transient to be studied in detail. It was clear that atomic oxygen was the most important negative ion emitted by the cathodes and its emission was therefore studied in some detail. Particular attention was paid to the oxygen ions that were emitted directly by the cathodes, rather than formed in the gas around them.

Changes in the oxygen ion current

During the activation of cathodes 8A, 8B and 8D, three similar series of curves were obtained of the energy distribution of the atomic oxygen ions. The most complete series of curves was for cathode 8B and it is shown in Fig. 4. The most rapid changes in the ion current occurred during the first hour.

The changes in the ion current are shown in another way in Fig. 5, where the logarithm of the oxygen ion current emitted by three cathodes 8B, 8C and 8F, is plotted against the time since their activation was begun. The corresponding

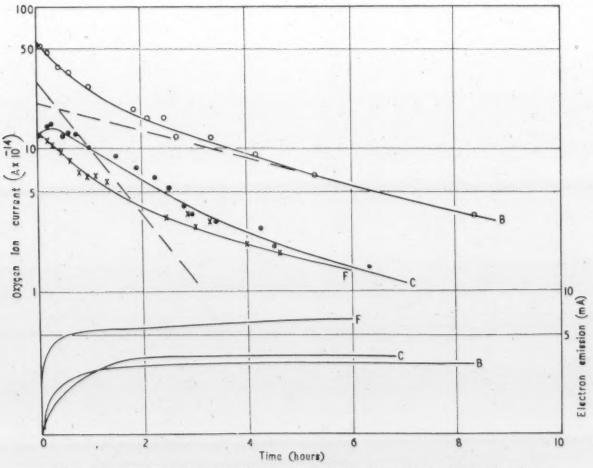


Fig. 5. Change of oxygen ion and electron currents during activation.
—×— oxygen ion current; F, electron current for cathode 8F at 1250° K.
—●— oxygen ion current; C, electron current for cathode 8C at 1350° K.
—○— oxygen ion current; B, electron current for cathode 8B at 1250° K.
(The ion current for 8B is scaled down by a factor of 10); the two components of the ion current from 8B are shown by dotted lines.

changes in electron current are also shown on the graph. The ion current fell while the electron current rose, and it continued to fall even after the electron current had become steady. In each experiment there was an unavoidable delay of a few minutes between the start of activation and first measurement of the oxygen ion current, and there may have been some rise of oxygen ion current during that interval. Cathode 8C showed such an initial rise of ion current for 16 minutes, and another cathode 9L showed a similar rise for 8 minutes.

The variation of the oxygen ion current I with time t could be expressed as the sum of two exponentially decreasing components, as in Eqn (1)

$$I = a \exp(-pt) + b \exp(-qt). \qquad (1)$$

In the equation a is greater than b, and p is greater than q, so that the first term is the more important when t is small, but the second term is dominant when t is large. The exponent q decreased by about 10% in 24 hours, and it also seemed to increase with the temperature of activation as shown in Table 1. The results for the first term were obtained from

Table 1. *Results for cathodes* 8B, 8C, 8F

Cathode	*Temperature* (° K)	*Parameters for Eqn* (1) $a(\times 10^{-14}$ A)	p (*per hour*)	$b(\times 10^{-14}$ A)	q (*per hour*)
8F	1250	7	1·3	5	0·23
8B	1300	300	1·2	210	0·25
8C	1350	11	1·1	3	0·29

Fig. 5 by subtracting the second term graphically from the total ion current, as shown for the curve for cathode 8B; consequently the results are less reliable for a and p than for b and q. Similar results were obtained with two nickel matrix cathodes containing triple carbonates, which were supplied by Dr. R. W. Fane of the English Electric Vacuum Research Unit.

As a working hypothesis it was assumed that the two terms of Eqn (1) represented two sources of emission of oxygen ions, and some experiments were devised to test this hypothesis. An attempt to measure how q varied with the temperature of a particular cathode did not give consistent results, because the cathode had to be run for so long before q was known with sufficient accuracy that other factors, such as changes in the concentration gradient of oxygen, had affected it more than the temperature. However, the experiment did show that both terms of Eqn (1) were increased when the temperature was raised.

Recovery from poisoning

Three cathodes were aged until the second term of Eqn (1) was dominant. Each was poisoned with oxygen by heating a filament coated with barium peroxide inside the mass spectrometer, and its recovery from poisoning was studied. The cathodes were repeatedly poisoned and allowed to recover at temperatures from 900 to 1300° K, modulator voltages 0–250 v, and degrees of poisoning from 10 to 90% reduction of electron emission. On recovery from poisoning, at temperatures below 900° K, the instrument was not sensitive enough to detect any oxygen ion emission. The cathodes always regained their original electron emission in a few minutes, and emitted a current of negative ions of atomic oxygen. The ion current decreased by a factor of 10 in 3 or 4 seconds, which was more than a thousand times faster than it had decreased during activation. The ion current was measured with the d.c. amplifier which drove a galvanometer of period 5·5 seconds, and such rapid changes could not be followed nor could any differences be detected that might have been caused by different conditions of poisoning. The mass spectrometer needed its anode voltage of 600 v left on throughout the experiments, so these results cannot be compared in detail with Richardson's (1957), which were performed at 8 v, but they confirm his general conclusions. Richardson's results for recovery from sulphur poisoning were similarly confirmed with two other cathodes; the recovery was similar to recovery from oxygen poisoning, but the current of negative ions ($^{32}S^-$) was smaller. The energy distribution of the oxygen ions just after recovery was the same as it had been just before poisoning, as shown in Fig. 6.

The effect of the modulator voltage

In the experiments on activation the modulator had been kept at 250 v positive with respect to the cathode. The modulator was 0·2 cm from the cathode and had a hole in it of only 0·1 cm diameter. It screened the cathode so effectively from the anode that it collected 99·8% of the electron current, and the activation of the cathode seemed to be unaffected by switching the anode voltage on and off. However, during the experiments on oxygen poisoning the modulator voltage was found to affect the energy distribution of the oxygen ions. This is illustrated by Fig. 6, which shows

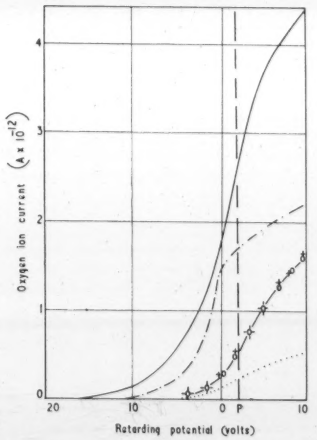


Fig. 6. Changes in energy distribution of oxygen ions:
——— cathode 8D, $V_1 = 250$ v.
– – – cathode 8D, $V_1 = 10$ v; + before poisoning, ○ after poisoning.
—·—· cathode 8F, $V_1 = 110$ v.
····· cathode 8F, $V_1 = 20$ v.

how a decrease in the modulator voltage from 250 v to 10 v decreased the proportion of high energy ions to low energy ions, and also the maximum energy of the ions. The precise maximum energy was uncertain, both because the ion current did not cut off sharply there, and because decreasing the modulator voltage decreased all the ion current, and so the current at the highest energy might have been undetected.

However, this result suggested that a study of the effect of the modulator voltage would give information about the type of positive ions that bombarded the cathode, and five cathodes were therefore activated at various values of modulator voltage.

When the potential difference V_1 between the modulator and the cathode was low the whole of the variation with time of the sputtered current of oxygen ions was described adequately by the second term of Eqn (1), but when V_1 was high the complete equation was needed to describe the variation. The activation of one cathode was successfully accomplished with $V_1 = 10$ v, but another cathode tested with $V_1 = 4$ v gave only a few microamperes of electron emission after three hours of activation. For both cathodes V_1 was raised to 100 v after three hours of activation and this caused an immediate increase in the emission of both electrons and oxygen ions. The electron emission did not continue to rise for the first cathode, but it did for the second one until it reached 7 mA after three more hours. The oxygen ion emission, which had previously been described by the second term of Eqn (1), fell rapidly after its sudden rise and continued to fall in the manner described by the complete equation. The third cathode was activated with V_1 alternately at 250 v and at 0 v for periods of thirty minutes at a time. The oxygen ion current fell rapidly like the first term of Eqn (1) when $V_1 = 250$ v, and slowly like the second term of the equation when $V_1 = 0$. Experiments with the two other cathodes showed that the oxygen ion current could be made to fall alternately rapidly and slowly by changing the modulator voltage between 15 v and 10 v. During the early stages of the activation of one of them the same effect was produced by switching V_1 between 13·75 v and 13·15 v, although later on 15 v was needed to restore the rapid fall of ion current. Changes of V_1 above 10 v had no perceptible effect on the change of oxygen ion emission described by the second term of Eqn (1). When V_1 was kept at 4 v the oxygen ion current was only just detectable and it was not certain what effect, if any, it had on the rate of change of the oxygen current.

The effect of the cathode's porosity

The second batch of cathodes had been prepared from a similar mixture and had been formed at various known pressures in a hydraulic press. This batch was studied at 1300° K and $V_1 = 250$ v, to discover whether either exponent p or q of Eqn (1) depended on the porosity of the cathode. After its activation each cathode pellet was removed from the mass spectrometer and its sintered porosity was calculated from its density. The parameters for Eqn (1) that fit the experimental results for each cathode are shown in Table 2.

The amount of oxygen emitted did not seem to depend on the porosity, but its rate of supply did. The exponent q was reproducible for pairs of cathodes formed at the same pressure, and for cathodes formed at high pressures q was approximately inversely proportional to the porosity. Cathode pellets formed at low pressures were so fragile that they usually split or flaked to pieces during the experiment, and that is why there is only one result for 10 ton/in^2 and only one result for a lower pressure.

4. Discussion

The primary source of the oxygen ions was the dissociation of barium oxide. Even when all other conditions were favourable for them they did not appear unless the cathode contained barium oxide, and therefore they were not formed from the residual gas in the mass spectrometer. Their energy distribution showed that they were removed from the cathode by positive ion bombardment, and the way in which their current decreased showed that they came from two secondary sources, which were responsible for the two terms in Eqn (1) and will be referred to as sources P and Q. The experiments on the effect of changing the potential difference between the modulator and the cathode showed that negative ions of oxygen were removed from Q at fairly low voltages, but were not removed from P until the potential difference was at least 13·75 v and generally not until it was 15 v. There was only about 0·5 v contact potential difference between cathode and modulator, presumably because barium oxide had evaporated from cathode to modulator during the heat treatment of the cathode. The result suggests that P had a high work function and a high adsorption energy for oxygen, whereas Q had low values of these energies.

It is suggested as a hypothesis that source Q is oxygen on the barium oxide at the ends of the pores in the cathode, and source P is oxygen on nickel remote from the ends of the pores. Both P and Q are initially supplied with oxygen by the dissociation of some barium oxide during the heat treatment of the cathode. When activation is begun at a high enough voltage then the electrons emitted by the cathode ionize the residual gas by collision and the positive ions so formed bombard the cathode with sufficient energy to remove oxygen as negative ions from both P and Q. A rough calculation shows that 10^{-6} A of electrons could have produced sufficient positive ions in the residual gas to sputter the observed current of negative ions from the cathode, if 0·1% of the cathode were covered with oxygen. A similar supply of positive ions could have been provided gif 10^{-11} of barium atoms had evaporated each second. The results of experiments on poisoning showed that sufficient positive

Table 2. *The effect of porosity*

Cathode	Forming pressure (ton/in^2)	Porosity (%)	Parameters for Eqn (1) $a(\times 10^{-14}$ A)	p (per hour)	$b(\times 10^{-14}$ A)	q (per hour)
9K	170	10	70	1·5	17	0·06
9L	170	10	100	1·9	31	0·05
9M	100	12·5	128	4·3	154	0·13
9N	100	12·5	*	*	60	0·11
9O	50	17	110	1·9	76	0·25
9P	50	17	48	2·3	40	0·22
10R	10	27·5	30	6·3	16	0·25
10U	4	33	50	2·7	13	0·22

* Results not obtained because of temporary fault in apparatus.

ions were formed to remove the oxygen very rapidly from the cathode, even when the electron emission current was quite small.

During recovery from poisoning the current of oxygen fell very rapidly because the oxygen that was removed from the cathode was not replaced. During activation the current of oxygen ions fell comparatively slowly because fresh oxygen was continuously being supplied by the dissociation of barium oxide. The fall in the oxygen ion current continued for at least 24 hours and so showed that the oxygen was being replaced less rapidly as time went by. Since the current of oxygen ions fell less rapidly from source Q than from source P, it follows that it was easier for fresh oxygen to reach Q than to reach P. The rate of fall of the current from Q increased with the porosity of the cathode, but that from P did not, showing that Q was supplied through the pores, but P was not. This result agrees with the hypothesis that Q was at the ends of the pores, but P was remote from them. Oxygen probably reached Q by Knudsen flow through the pores, but reached P by surface diffusion over the crystals.

Acknowledgments

The author thanks Prof. F. A. Vick and Prof. D. J. E. Ingram for laboratory facilities during 1959 and 1960 respectively, Dr. J. F. Richardson and Dr. R. W. Fane for the gift of materials; he also thanks them, and his colleagues, for discussion of this work.

References

BALAS, W., DEMPSEY, J., and REVER, E. F., 1955, *J. Appl. Phys.*, **26**, 1163.

BECK, A. H., BRISBANE, A. D., CUTTING, A. D., and KING, G., 1954, *Vide, Paris*, **9**, 302.

FANE, R. W., 1958, *Brit. J. Appl. Phys.*, **9**, 149.

HAMAKER, H. C., BRUINING, H., and ATEN, A. H. W., 1947, *Philips Res. Rep.*, **2**, 171.

MACNAIR, D., LYNCH, R. T., and HANNAY, N. B., 1953, *J. Appl. Phys.*, **24**, 1335.

RICHARDSON, J. F., 1957, *Brit. J. Appl. Phys.*, **8**, 361.

RICHARDSON, J. F., and VICK, F. A., 1960, *Brit. J. Appl. Phys.*, **11**, 73.

SURPLICE, N. A., 1960, *Brit. J. Appl. Phys.*, **11**, 430.

VICK, F. A., and WALLEY, C. A., 1954, *Proc. Phys. Soc. B*, **67**, 169.

Emission of negative ions of oxygen from dispenser cathodes Part 2.—Cathodes of barium aluminate in sintered tungsten

by N. A. SURPLICE, B.Sc., Ph.D., A.Inst.P., Physics Department, University College of North Staffordshire, Keele, Staffs.

MS. received 14th November 1960

Abstract

A cathode of sintered tungsten impregnated with barium aluminate has been studied in a simple mass spectrometer and has been found to emit negative ions of oxygen. About 90% of the oxygen ions were in the atomic form and 10% in the molecular form. The experiments have shown that the ions come from an adsorbed layer of oxygen on the tungsten matrix, from which they are removed by bombardment with positive ions, chiefly of nitrogen. Similar results were obtained with a tungsten matrix without any impregnant.

1. Introduction

THE impregnated type of dispenser cathode consists of a porous tungsten matrix impregnated with barium aluminate (and sometimes a little calcium oxide), rather than with the mixture of barium and strontium oxides that is used in the ordinary oxide-coated cathode. It has been described in detail by Levi (1953, 1955, 1957/58). Both its mechanism of operation and the rate at which barium evaporates from it have been studied by Rittner *et al.* (1957, 1958) and by Brodie *et al.* (1957, 1959). The barium is generated by a chemical reaction between the impregnant and the matrix, then diffuses over the tungsten and thus reduces its work function. The evolution of oxygen is not an essential part of the cathode's mechanism. However, the tungsten would be expected to be covered with a chemisorbed layer of oxygen from the atmosphere, and the experiments described below show that some of the oxygen is removed as negative ions when the cathode is in use.

2. Apparatus

The barium aluminate impregnated cathode was supplied by Dr. A. G. Mitchell of the English Electric Valve Co., and a similar tungsten cylinder without any impregnant was supplied later on request. The cathode was in the form of a cylinder 0·63 cm long open at both ends, it was of 0·55 cm external diameter and 0·43 cm internal diameter, and the tungsten matrix was 25% porous. The apparatus that was used to study it has already been described in Part 1 of this paper (Surplice 1961).

The cathode was made the ion source of the mass spectrometer and its temperature was measured with a tungsten–nickel thermocouple welded to its outer surface. The mass spectrometer was outgassed at 400° C for 18 hours. The cathode was kept at 1250° K while the electron gun was eddy-current heated and while the instrument was sealed off and gettered.

The tungsten cylinder was studied both with the ordinary mass spectrometer and with a modified one. In the modified instrument a small oxide-coated cathode was mounted in a side tube on the glass envelope, so that it was clear of the electron gun, but could emit a current of electrons to the anode and thus ionize the residual gas there.

3. Results

The mass spectrum

The mass spectrum of the impregnated cathode contained negative ions of chlorine and oxygen. The peak ion currents to the collector, when it was a few volts positive with respect to the cathode, were of the order of 5×10^{-13} A of atomic oxygen and of chlorine, and about a tenth as much molecular oxygen. The energy distribution of each ion was found by the retarding potential method (Fig. 1). The results showed

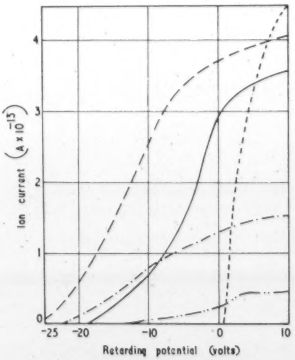


Fig. 1. Energy distributions:

ion from impregnated cathode — $^{16}O^-$; $^{32}O^-$; $^{35}Cl^-$

from cold tungsten cylinder — $^{16}O^-$

from hot tungsten cylinder — $^{16}O^-$

that more than half of the atomic and of the molecular ions of oxygen were sputtered off the cathode, but no chlorine ions had been emitted directly from the cathode. The chlorine probably came either from the glass or from accidental handling of the cathode. The energy distributions did not show any systematic changes with time.

The tungsten matrix without the impregnant could not be heated sufficiently to make it emit electrons, and it showed no negative ion emission in the ordinary mass spectrometer.

However, when it was put in the modified instrument and the auxiliary cathode was used to ionize the residual gas, then the mass spectrum was qualitatively similar to that of the impregnated cathode. The current of molecular oxygen ions was smaller from the tunsten matrix than from the impregnated cathode. Atomic oxygen ions of high energy appeared even when the tungsten was cold, but twice as many appeared when it was hot, so they probably came half from the tungsten and half from the auxiliary oxide-coated cathode.

After the experiments, the cathode was found to be pitted on the side opposite the 1 mm × 3 mm slot in the modulator, which suggested that it had been struck by high-energy positive ions from the anode.

Changes in the atomic oxygen ion current.

The part of the atomic oxygen ion current that was sputtered from the cathode changed with time in a manner that depended markedly on the potentials of the modulator and anode of the electron gun. It fell rapidly unless the modulator potential was less than 15 v positive to the cathode and the anode voltage of 700 v was switched off. The current of oxygen ions sputtered from the tungsten matrix fell in a similar manner, as shown in Fig. 2. The current was not any simple function of time.

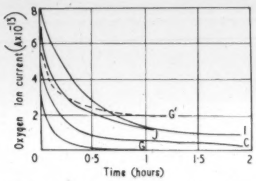


Fig. 2. Changes of oxygen ion current when voltage was left on:
ions from impregnated cathode { C, G: V_a = 580 v, V_1 = 250 v. I, J: V_a = 610 v, V_1 = 7 v.
ions from tungsten cylinder G′: V_a = 600 v, V_1 = 250 v.

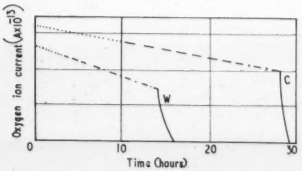


Fig. 3. Changes of oxygen ion current when anode voltage was left off:
C, impregnated cathode, *W*, tungsten cylinder.
..... V_1 = 7·0 v. — · — · V_1 = 14·4 v.
– – – – V_1 = 12·75 v. ——— V_1 = 15·0 v.

Both the impregnated cathode and the tungsten matrix were studied with the modulator voltage at selected values near the ionization potentials of the residual gas, and the anode voltage switched off except when measuring the ion current. The oxygen ion current fell only slowly while the modulator voltage V_m was increased step by step past the ionization potentials of barium (5·2 v), oxygen (13·6 v) and oxides of carbon (14·1 and 14·4 v) but it fell rapidly when V_m was raised above the ionization potential of nitrogen (14·55 v). These results are shown in Fig. 3.

4. Discussion

The source of the oxygen ions seems to have been a chemisorbed layer of oxygen on the tungsten, which was sputtered off mainly by nitrogen positive ions. Since 6·5 ev would be enough to remove the adsorbed oxygen, and 4·5 ev enough to give it an electron from the tungsten, it was surprising that it was not removed by oxygen (which would have been ionized when there was 13·6 v between the modulator and cathode, and have reached the cathode from the region of the modulator with a maximum energy of 13·6 ev). The oxygen appears to have been on tungsten, not on barium-on-tungsten in the impregnated cathode, as otherwise the effect of the modulator voltage would have been apparent at a lower voltage for the impregnated cathode than for the unimpregnated tungsten. Although the auxiliary oxide-coated cathode would also emit oxygen ions from the continual dissociation of barium oxide, its ions would only account for the ones detected when the tungsten cylinder was cold, and not when it was hot.

Acknowledgments

The author thanks Professor D. J. E. Ingram for laboratory facilities and Dr. A. G. Mitchell for the gift of cathodes.

References

BRODIE, I., and JENKINS, R. O., 1957, *J. Electronics and Control*, **2**, 457.

BRODIE, I., JENKINS, R. O., and TRODDEN, W. G., 1959, *J. Electronics and Control*, **6**, 149.

—— 1958, *J. Appl. Phys.*, **29**, 834.

LEVI, R., 1953, *J. Appl. Phys.* **24**, 233.

—— 1955, *J. Appl. Phys.*, **26**, 639.

—— 1957/58, *Philips Tech. Rev.*, **19**, (6), 186.

RITTNER, E. S., RUTLEDGE, W. C., and AHLERT, R. H., 1957, *J. Appl. Phys.*, **28**, 1468.

SURPLICE, N. A., 1961, *Brit. J. Appl. Phys.*, **12**, 214.

Some aspects of the self-heating and ignition of solid cellulosic materials

by P. H. THOMAS, Ph.D., M.A., G.I.Mech.E., and P. C. BOWES, B.Sc., A.R.C.S., A.R.I.C., Department of Scientific and Industrial Research and Fire Offices' Committee Joint Fire Research Organization, Boreham Wood, Hertfordshire

MS. received 5th May, in revised form 24th October 1960

Abstract

The self-heating and self-ignition of fibre insulating board and similar materials are discussed in the light of the well-known theory of thermal explosion. It is pointed out that the simple theory for a single reaction neglecting reactant loss during the induction period is inadequate for the interpretation of all the data published by Mitchell (1951). *In the case of wood fibre insulating board it is suggested that the reaction leading to significant increases in temperature at relatively low ambient temperatures is not that responsible for ignition; ignition appears to be mainly the result of other reactions which become appreciable only at higher temperatures. Some experimental results in support of this contention are given, and it is estimated that the total heat of reaction is* 80 *cal* g^{-1}.

The induction times reported by Mitchell for three materials show one of them (cotton linters) to be different in its behaviour from the other two.

Introduction

THE mathematical theory of thermal ignition (Semenoff 1928, Todes *et al.* 1933, 1939, 1940, Rice, Allen and Campbell 1935, Frank–Kamenetskii 1939) was developed initially for certain gas mixtures in which ignition or explosion depends primarily on the balance between the heat produced by the reaction and the heat lost to the surroundings. This theory has found application to the initiation of explosions in solids (Rideal and Robertson 1948, Bowden and Yoffe 1952, Ubbelohde and Woodward 1948, Cook 1958), and it is of interest to consider its application to the much slower ignition of solids such as wood, in which relatively slow exothermic reactions may play an important part under certain conditions (Gross and Robertson 1958, Akita 1956). On the basis of such theory one can, in principle, obtain the chemical and physical constants that are necessary to predict the critical sizes of pieces of material that would self-heat to ignition at any given temperature. Even if self-heating is not the primary practical interest, many problems in the ignition and combustion of these materials require the values of the chemical kinetic constants to be known.

Published results on the self-heating and ignition of fibre-insulating boards (Mitchell 1951, Gross and Robertson 1958, Akita 1956) will be discussed here in the light of thermal ignition theory which is based in the first place on the conventional assumption of very simple kinetics. Application of this theory then makes it clear that more complicated kinetics must be postulated to reconcile the results of both self-heating and ignition experiments; the simple theory, nevertheless, is useful for limited purposes.

Theory of thermal explosions

In this section we summarize some of the results of the general theory of thermal explosions.

The differential equation for the temperature distribution in a conducting solid in which heat is generated is

$$\nabla^2 T = \frac{1}{k}\frac{dT}{dt} - \frac{q(T)}{K} \tag{1}$$

where ∇^2 is the Laplacian operator, T is the absolute temperature, k is the thermal diffusivity, $q(T)$ is the rate of generation of heat per unit volume of the solid at T, K is the thermal conductivity and t is the time.

We assume initially that heating is due to a single reaction of nth order and that the Arrhenius law is obeyed. We shall also assume that the reaction is not limited by air supply.

We can then write

$$q = -\frac{d\lambda}{dt}Q \tag{2a}$$

and

$$\frac{d\lambda}{dt} = -f\lambda^n \exp(-E/RT) \tag{2b}$$

where Q is the heat of reaction per unit mass of reactant, λ is the concentration of mass of reactant per unit volume of total solid, f is a constant, E is the activation energy (cal/mol) of the reaction, and R is the universal gas constant.

For a solid which is heated symmetrically we have

$$\nabla^2 = \frac{d^2}{dx^2} + \frac{P}{x}\frac{d}{dx}$$

where x is a distance coordinate and $P = 0$ for slab, $P = 1$ for radial flow in a cylinder, and $P = 2$ for radial flow in a sphere.

Initially at $t = 0$, we have $\lambda = \lambda_0$ and $T = T_A$ the ambient temperature.

At the surface, the boundary condition is taken as

$$H(T - T_A) = -K\frac{dT}{dx}, \quad (x = \pm r) \tag{2c}$$

where H is the heat transfer coefficient at the surface and r is the half width of the slab or the radius of the cylinder or sphere.

The second boundary condition is by symmetry

$$\frac{dT}{dx} = 0 \text{ at } x = 0.$$

We employ the dimensionless quantities

$$\delta = \frac{QfEr^2\lambda_0^n}{KRT_A^2}\exp(-E/RT_A) \tag{3a}$$

$$\theta = \frac{E(T - T_A)}{RT_A^2} \tag{3b}$$

$$\left.\begin{aligned} \alpha &= Hr/K & z &= x/r \\ \tau &= kt/r^2 & B &= \frac{Q}{\rho c}\frac{E\lambda_0}{RT_A^2} \end{aligned}\right\} \tag{3c}$$

where ρ is the density of the solid material and c the specific heat, from which we can write Eqns (1) and (2b) as

$$\nabla^2\theta = \frac{d\theta}{d\tau} - \delta e^{\theta}\left(\frac{\lambda}{\lambda_0}\right)^n \exp\left(-\frac{RT_A^2\theta^2}{ET}\right) \tag{4a}$$

(∇^2 is here taken with respect to z),

and
$$\frac{1}{\lambda_0}\cdot\frac{d\lambda}{d\tau} = -\frac{\delta}{B}e^{\theta}\left(\frac{\lambda}{\lambda_0}\right)^n \exp\left(-\frac{RT_A^2\theta^2}{ET}\right). \quad (4b)$$

Now if $T \sim T_A$ and θ is of order unity (Semenoff 1928, Todes *et al.* 1933, 1939, 1940, Rice, Allen and Campbell 1935, Frank–Kamenetskii 1939), and if $E/RT \gg 1$, we can treat the exponential in θ^2 as unity. After writing λ/λ_0 as w we obtain

$$\nabla^2\theta = \frac{d\theta}{d\tau} - \delta w^n e^{\theta} \quad (5a)$$

$$\frac{dw}{d\tau} = -\frac{\delta}{B}w^n e^{\theta}. \quad (5b)$$

Since the only new quantity in the boundary and initial conditions is α the solutions of these equations are of the form

$$\left.\begin{aligned}\theta &= \theta(\delta, B, n, \alpha, \tau, z)\\ w &= w(\delta, B, n, \alpha, \tau, z)\end{aligned}\right\}. \quad (6)$$

Curves of θ_0 (the centre temperature) against τ are of two kinds (Semenoff 1928, Todes *et al.* 1933, 1939, 1940, Rice, Allen and Campbell 1935). For low values of δ, θ_0 increases with τ to a maximum of order unity and then falls. For large values of δ and provided the dimensionless adiabatic temperature rise, B, is large compared with unity, θ_0 plotted against τ passes through an inflection when it is of order unity and then rises rapidly to a maximum of order B. This type of temperature behaviour is conventionally considered as characteristic of ignition, and the lowest value of δ for which a curve of θ_0 against τ includes such an inflection while the temperature is rising is then defined, again conventionally, as the critical value of δ for ignition δ_c and this must be given functionally by

$$\delta_c = \delta_c(B, \alpha, n). \quad (7a)$$

At this inflection, or ignition point, for any δ greater than δ_c,

$$\left.\begin{aligned}w_i &= w_i(B, \alpha, n, \delta)\\ \tau_i &= \tau_i(B, \alpha, n, \delta)\\ \theta_i &= \theta_i(B, \alpha, n, \delta)\end{aligned}\right\}. \quad (7b)$$

The values of w_i, θ_i and τ_i when 'ignition' just occurs ($\delta = \delta_c$) are functions of B, α and n only and are independent of the size of the specimen or its ignition temperature, except for the secondary effect of T_A on B and θ.

General analytical solutions are not available, but certain special cases have been solved. Thus, when n is zero or when B is effectively infinite, Eqn (5*b*) is not relevant to solving equation (5*a*) nor to determining the critical values of δ_c, θ_i or τ_i, which are each then functions of α only. The well-known stationary state solutions for infinite α given by Frank–Kamenetskii (1939) then apply. Thomas (1958) has given solutions for all values of α for these limiting values of n and B, whilst Kinbara and Akita (1960) have devised an approximate method of solving the transient equation for an infinite value of B/n.

Approximate analytic solutions to equations (5*a*) and (5*b*) for symmetrical heating, with finite values of B/n and $\alpha \to 0$ have recently been obtained (Thomas, to be published). These are very similar to those calculated by the numerical integration method of Rice, Allen and Campbell (1935) for $B > 20$, and show that the effect of B on δ_c is given (Thomas, to be published) asymptotically by

$$\delta_c(B) \simeq \delta_c\left\{1 + 3\left(\frac{n}{B}\right)^{\frac{2}{3}}\right\}. \quad (8)$$

When B is small it may still be possible to distinguish two types of curves for θ_0 plotted against τ as described above, but this mathematical distinction begins to lose physical meaning. The detailed calculation by Todes and Melentiev (1935, Fig. (*f*), p. 1606) for $B = 9$ shows that there ceases to be a sharp distinction between ignition and non-ignition for a value of B as low as this and it is then difficult to ascribe a critical value to δ.

If we use Eqn (8) to obtain a first approximation to $\delta_c(B)$ with B equal to 9, we find δ_c is raised by less than a factor of 2 and hence it is reasonable to assume that, if there is a sharply defined ignition, the value of δ_c is not more than about twice the value calculated for no loss of reactant. Buben (1945) has analysed the effect of the diffusion of a reactant on the critical value of δ and, for this case too, the existence of a sharply defined reaction implies a value of δ_c not more than about twice the value calculated for the case of infinite B/n.

Analysis of ignition data

We shall use the theory outlined above to interpret Mitchell's ignition data for fibre insulating board and similar materials (Mitchell 1951) which self-heat as a result of exothermic decomposition.

It follows from the definition of δ, that a plot of $\ln(\delta_c T_A^2/r^2)$ against $1/T_A$, where T_A is the 'ignition temperature' of a specimen of size r, should be linear with a slope of $-E/R$.

The results of Mitchell (converted to c.g.s. units and degrees centigrade) for the ignition of piles of wood fibre insulating board in an oven, and some recent results obtained by us for both cubes and vertically hung square slabs of wood fibre insulating board are plotted in the above way in Fig. 1(*a*). Fig. 1(*b*) shows Mitchell's results for piles of cane fibre insulating board and cotton linters. As a first approximation the heat of reaction, and therefore B, was initially assumed to be large enough for B not to influence δ, so that theoretical values of δ_c as a function of α only could be used.

The theoretical value of δ_c for a pile of octagonal section,

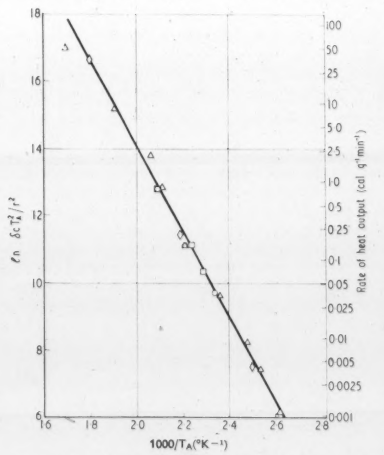


Fig. 1(*a*). Correlation of ignition data for wood fibre insulating board.

△ Mitchell (1951).
○ cube, □ slab } Authors.
◇ Gross and Robertson (1958) (calculated rate of heat evolution). $E = 25\,000$ cal mol^{-1}.

such as was actually used by Mitchell, may be estimated using the approximate method described by Frank–Kamenetskii (1955). Thus for large values of B/n and α, the value of δ_c for a cylinder of length equal to the diameter is found to be approximately 2·8 compared with 3·32 for the sphere. For a cube it is approximately 2·5. For an octagonal pile we may therefore conveniently take $\delta_c = 2{\cdot}65$, and the values for a sphere for finite values of α reduced accordingly by 20%, have been used.

The cooling coefficient was calculated assuming the surface of the specimen to be black and allowance was made for natural convection. The calculated values of α and δ_c for a sphere are shown in Table 1 for the specimens of wood fibre

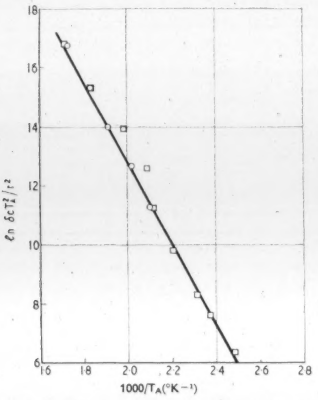


Fig. 1(*b*). Correlation of ignition data (Mitchell).
○ Cotton linters; □ Cane fibre insulating board.

insulating board of the various sizes used by Mitchell. r was taken as the half-width of the piles. The effect on δ_c of the variation of α is only significant for the smallest specimens. For the vertically hung slabs α was calculated in the same way and the appropriate correction made to δ_c.

Table 1. *Calculated dimensionless heat transfer coefficient α and corresponding value of δ_c for Mitchell's data (B/n assumed infinite)*

Size of pile r (in.)	α	δ_c *sphere*	*Estimated* δ_c *for octagonal pile*
$\frac{1}{16}$	2·5	1·7	1·4
$\frac{1}{8}$	3·0	1·9	1·5
$\frac{1}{4}$	4·4	2·2	1·8
$\frac{1}{2}$	7·2	2·6	2·1
1	11·6	2·8	2·2
2	17·3	2·9	2·3
4	30·0	3·1	2·5
6	41·0	3·2	2·6
11	67·0	3·2	2·6

The data do, in fact, appear as straight lines. The activation energies obtained from the slopes of these lines are given in Table 2 together with those obtained by Gross and Robertson (1958), who allowed material to self-heat in an adiabatic calorimeter (Raskin and Robertson 1954).

For large values of α the difference between the surface temperature and the ambient temperature tends to zero and δ_c becomes independent of α. It follows from Eqn (7*a*) that δ_c then depends only on B and n, i.e. on the nature of the reaction, and data for a given material can therefore be extrapolated to large values of r without assuming a particular value for δ_c. This condition is fulfilled for the larger specimens listed in Table 1, and the deviation for the smaller specimens is in fact negligible when the points are plotted in Fig. 1. Any prediction of ignition temperatures for large specimens by extrapolation then depends only on the requirement that the effective value of E is constant over the desired range of temperature.

Table 2. *Values of activation energy (cal/mol)*

Material	*Derived from Mitchell's ignition data*	*Measured by Gross and Robertson (1958)*
Wood fibre insulating board	25 000	25 700
Cane fibre insulating board	28 000	—
Cotton felted linters	28 000	34 500

If we assume that the values of δ_c quoted in Table 1 are applicable, the value of $(Q\lambda_0^n f e^{-E/RT})/\rho$, i.e. $q(T)/\rho$, at any temperature may be found from ignition data. Then, since the rate of heat generation per gram of total substance is, from the definition of $q(T)$ and δ_c,

$$\frac{q(T)}{\rho} = \frac{KR}{E}\frac{\delta_c}{\rho}\frac{T^2}{r^2} \tag{9}$$

the ordinate in Fig. 1 may be readily converted into a heat generation scale; this has been done using the data for wood fibre insulating board in Fig. 1(*a*).

Gross and Robertson (1958) have measured rates of heat output, and three points calculated from their data are also shown in Fig. 1(*a*). The agreement between theoretical and measured values is satisfactory for, despite the difference in the numerical values of E in Table 2, the effect of this on calculated ignition temperatures is small. However, in the case of cotton linters there is a tendency for the ignition temperatures predicted from their calorimetric data to be increasingly greater than the values obtained by Mitchell as the size of the specimen is increased.

Analysis of self-heating data

Provided δ does not exceed δ_c a steady state can theoretically exist until reactant loss becomes the controlling factor.

For the non-critical steady state it is possible to calculate the relation between θ_0 and the value of δ if reactant loss is neglected. This computation has been performed for certain values of α assuming no loss of reactant (Appendix 1) and the results are shown in Fig. 2.

In his self-heating experiments Mitchell measured the central temperature rises above ambient, i.e. $T_0 - T_A$, for various values of the ambient temperature T_A and various sizes of specimen. The relation between δ and θ_0 in Fig. 2 could be used to determine E from Mitchell's data, but the procedure would be somewhat cumbersome if the whole range of experimental values of $T_0 - T_A$ were to be included. If, however, only small degrees of self-heating are considered, a simplified procedure is practicable as follows.

It may be readily shown that if, on the right-hand side of the steady-state version of Eqn (5*a*), e^{θ} is approximated to $\exp\theta_0$ (a constant); the corresponding approximate solution for a sphere, neglecting loss of reactant, is

$$\theta_0 = \frac{\delta}{6}\frac{2+\alpha}{\alpha}\exp\theta_0. \tag{10}$$

δ as a function of θ_0 according to this equation is also shown in Fig. 2 for various values of α and it may be seen that for $\theta_0 < 0{\cdot}4$ this equation is not an unreasonable approximation.

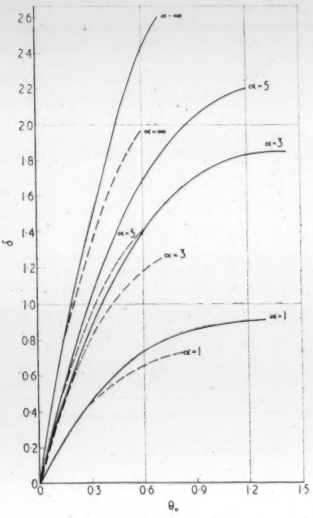


Fig. 2. Non-critical values of δ and θ_0 for various cooling conditions for a sphere.
Dotted lines represent $\delta = 6\theta_0 \exp(-\theta_0)\alpha/(2+\alpha)$.

If θ_0 is to be less than $0{\cdot}4$ then $T_0 - T_A$ must be less than $0{\cdot}4(T_A{}^2R)/E$, which for $T_A = 600°$ K, $R = 2$ cal mol^{-1} deg C^{-1} and E approximately 25 000 cal/mol, gives a maximum allowable value of $T_0 - T_A$ equal to 12° C.

If $T_0 - T_A \ll T_A$, $\delta \exp \theta_0$ is approximately equal to $\{Q\lambda_0{}^n f E r^2 \exp(-E/RT_0)\}/KRT_A{}^2$.

It may be estimated that the temperature rise in an octagonal pile (as used by Mitchell) is about 50% greater than in a sphere of diameter equal to the width between the parallel faces of the octagonal pile. We therefore replace the factor 6 by 4 in Eqn (10). It then follows from Eqn (10), the definitions of θ and δ, and the above approximation for e^θ, that

$$\frac{\alpha}{2+\alpha}\frac{T_0 - T_A}{r^2} \simeq \frac{Q\lambda_0{}^n f}{4K}\exp(-E/RT_0) \qquad (11)$$

i.e.

$$\frac{q(T_0)}{\rho} \simeq \frac{4K}{\rho}\frac{\alpha}{2+\alpha}\frac{T_0 - T_A}{r^2}. \qquad (12)$$

A plot of

$$\ln\left(\frac{\alpha}{2+\alpha}\frac{T_0 - T_A}{r^2}\right)$$

should therefore vary linearly with $1/T_0$ and this is shown in Fig. 3. The ordinate can be related by Eqn (12) to the rate of heat generation per unit mass, $q(T)/\rho$, in the same way as can the ignition data. Here, however, no theoretical value need be assumed for δ_c.

The curvature in Fig. 3 is greater than can be accounted for by the approximation in Eqn (10). Nevertheless, in the region where the temperature differences are small the curves are asymptotic to a common line which corresponds to an activation energy of 21 500 cal/mol. This is somewhat lower than the value given by ignition data.

Much more significant is the estimate of heat output derived from these experiments for small amounts of self-heating. It is in fact an order greater than found from measurements by Gross and Robertson in the adiabatic calorimeter or from Mitchell's experimental data on ignition. Thus at a value of $1000/T°$ K equal to $2{\cdot}3$ ($T = 161°$ C), the ignition data of Mitchell in Fig. 1 give a value of about $0{\cdot}07$ cal g^{-1} min^{-1} for the heat output, and the self-heating data in Fig. 3 a value of $0{\cdot}70$ cal g^{-1} min^{-1}. (It should be noted that the ratio of the rates of heat generation given in Eqns (9) and (12) is independent of K and ρ.)

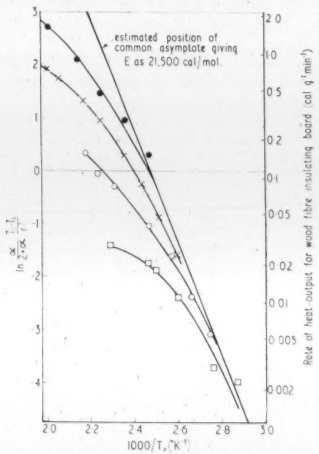

estimated position of common asymptote giving E as 21,500 cal/mol.
1000/T, (°K⁻¹)
Rate of heat-output for wood fibre insulating board (cal g⁻¹min⁻¹)

Fig. 3. Correlation for Mitchell's self-heating data.
● 1 in. block ($2r$); ○ 4 in. block;
× 2 in. block; □ 8 in. block.

The same point can be made another way. The maximum temperature rise ΔT in a sphere of radius r and conductivity K in which heat is generated at a uniform rate q may be calculated from conduction theory (Carslaw and Jaeger 1947, p. 207) as $\Delta T = qr^2/6K$ and for an octagonal pile it may be estimated as approximately 50% greater than this. Now at 161° C the kinetic constants given by Gross and Robertson give a value for q of $0{\cdot}25 \times 10^{-3}$ cal cm^{-3} sec^{-1} ($0{\cdot}055$ cal g^{-1} min^{-1}). Thus for r equal to $1{\cdot}27$ cm and K equal to $1{\cdot}2 \times 10^{-4}$ c.g.s. units, ΔT would be approximately 1° C; a result which justifies assuming a uniformly distributed rate of heating. For such a pile at 161° C Mitchell recorded a rise of nearly 9° C.

The above theory has assumed the existence of a steady state. This is not so but, since loss of reactant would lower the observed temperature rise, the difference between the two sets of data is actually underestimated.

D. Gross has communicated to us privately two complete temperature-time curves for 2 in. and 4 in. octagonal prisms of wood fibre insulating board obtained at the National Bureau of Standards, Washington. These were for heating in an oven at 133° C. From the rates of temperature rise of the specimens when their temperatures equalled the oven

temperature, the rates of heat output at 133° C can be calculated.

Assuming a specific heat of 0·3 cal g^{-1} deg C^{-1}, these rates are 0·12 and 0·09 cal g^{-1} min^{-1}, which agree to within 20% with the result calculated from the assymptotic line in Fig. 3.*

Thus the rates of heat output calculated in two ways from self-heating experiments are consistent, but larger than the values calculated from ignition data.

We conclude that there is a discrepancy between self-heating data and ignition data for wood fibre insulating board when compared on the assumption of simple reaction kinetics; this is discussed below.

Effect of more than one reaction

It is clear from the results of differential thermal analysis that the decomposition reactions for cellulose and lignin are complicated (Berkowitz 1957, Breger and Whitehead 1951) and therefore the treatment of self-heating and ignition in terms of a single reaction may be too drastic a simplification.

Recently some experiments have been made to observe the temperature/time relation for a specimen heated in an oven for a long time. The result shown in Fig. 4 is for a 1 in. cube of wood fibre insulating board heated at 175° C; it has been corrected for the small difference between the temperature of the specimen and the temperature of the oven atmosphere, that was a consequence of the oven ventilation. The graph shows a maximum temperature difference between the centre of the specimen and the oven occurring within a few minutes after the specimen has reached the oven temperature (taken as time zero), followed by a long period during which the temperature difference falls gradually to zero.

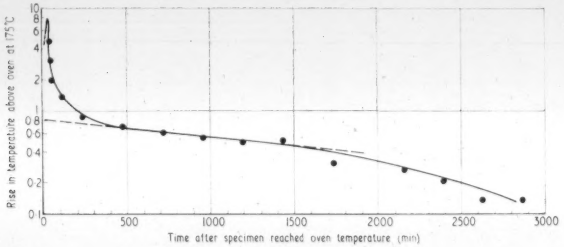


Fig. 4. Self-heating of 1 in. cube of wood fibre insulating board at 175° C.

Now, although the reaction rate varies with temperature, the temperature difference between the specimen and the oven is small and we can base an approximate theory on a rate of reaction which is independent of the temperature. Thus, in Eqns (5*a*) and (5*b*) we write e^{θ} equal to a where a is a constant approximately equal to unity. The solution of the equation is then readily obtained, using products of Fourier series (Carslaw and Jaeger 1947, Chap. 6), to give the central temperature for the cube as a function of time, α, δ and ^{0}B (see Appendix 2). According to this calculation there should be an approximately exponential decrease after 15 min if the reaction is first order.

The curve in Fig. 4, however, is not exponential and shows an inflection over the period 500 to 1500 min. This cannot be accounted for by a reaction of constant order; further, if the portion of the curve beyond 1500 min is neglected (thereby eliminating the inflection), calculations show that the curve cannot even then be explained by a single reaction of constant order. It seems therefore that more than one reaction is involved.

It can be shown (see Appendix 2) that, provided the reactions approach completion and the rise in temperature $T - T_0$ is not too large, the area under the temperature–time curve is related simply to the total heat produced, whatever the manner in which it is released as a function of time, i.e.

$$\int_0^{\infty} (T - T_0)dt = \frac{MQr^2\lambda_0}{K} \qquad (13)$$

where M is a constant for any given value of α.

For a 1 in. cube with a surface heat transfer coefficient of 6×10^{-4} c.g.s. units and $K = 1{\cdot}2 \times 10^{-4}$ c.g.s. units, $\alpha \simeq 6$ and M is 0·29; for infinite α it is 0·24.

The value of $Q\lambda_0/\rho$, i.e. the heat of reaction per unit mass of total substance obtained by applying Eqn (13) to Fig. 4, is about 80 cal/g of total substance, of which only about 16 cal/g is liberated in the initial rise and fall to a temperature of 0·8° C above oven temperature (see Fig. 4) and is attributable to an initial short-lived reaction. This value of 80 cal/g supersedes the earlier estimated value of 150 cal/g based on uncorrected data (Thomas 1959). Clearly the long tail in Fig. 4 represents the important source of heating because the total B for both reactions is about 13, while for the first reaction alone it is only about 3 which, as has been shown above, is too small to produce the sharp ignition observed in practice. It follows that in wood fibre insulating board there is a fast reaction of low heat output which is not primarily responsible for the observed ignition; the slow reaction that becomes important at higher temperatures, is the one responsible for ignition of the specimen.

Previous estimates of the heat of reaction ($Q\lambda_0/\rho$ above) reviewed by Byram *et al.* (1952), range between 50 and 100 cal/g for wood. Akita (1956), more recently, has given a value of 32 cal/g for wood and this value has been used by Kinbara and Akita (1960) in discussing transient problems. However, this gives a value for B of 6, and suggests it is too low to reconcile with sharp ignition.

It is to be noted that the maximum temperature rise in Fig. 4 is 8° C, while the maximum temperature rise due to the second reaction above is about 0·8° C—these values are in approximately the same ratio as those in the self-heating and ignition data of Mitchell.

This picture of two exothermic reactions leads to the conclusion that comparison of *ignition temperatures* for various

* These rates are relatively somewhat higher than those that may be calculated from rates of temperature rise at 325° F (163° C) recently reported by Smith (1959) for wood fibre insulating board in an adiabatic calorimeter. Smith shows that self-heating is related to the amount of components in the wood extractable with solvent.

sizes of specimen will give an estimate of the activation energy for the second reaction, whereas the measurement of the maximum *self-heating* temperature tends to give that for the first reaction. If the difference in the activation energies determined for the two sets of data given by Mitchell is real, one would conclude that the first reaction has a slightly lower activation energy than the second—which is consistent with its being relatively more important at lower temperatures.

On this view the difference between the straight asymptotic line in Fig. 3 and the curved experimental lines reflects the loss of reactant for the first reaction. Some of this loss may occur in the heating up to the oven temperature, in which case the amount of loss will be dependent on the size of the specimen because this affects the rate of heating up to the oven temperature.

Equation (8) shows that a value of B of 13 raises the ignition parameter δ_c for infinite B/n by about 50% and this would lead to a similar increase in the estimate of the heat output in Fig. 1(*a*)—this heat output scale having been derived on the assumption of no loss of reactant. The data obtained by Gross and Robertson gives the actual heat output, which is less than the value assuming no loss of reactant and therefore both sets of data should be corrected in the same direction. Although the corrections are of different magnitude their effect on the correlation of the two sets of data in Fig. 1(*a*) is small and the corrections have not actually been made.

Interpreting the result in Fig. 4 as evidence for two exothermic reactions one would expect that a short preheating of a specimen of wood fibre insulating board at a temperature below its ignition temperature would lead to exhaustion of the first reaction and a consequent slight increase in ignition temperature. This has in fact been found; thus, the ignition temperature of a 1 in. cube is raised about 10° C after preheating for about ½ h at 170° C, but it has not yet been possible to relate this increase simply and quantitatively to the foregoing analysis. It is not at all unlikely that our picture of the exothermic processes is still over-simplified.

Induction time

The induction time is a fundamental part of any discussion of self-ignition.

Zinn and Mader (1960) have recently performed numerical calculations for the induction times t_i of spheres when the initial temperature of the material is below the ambient temperature in an ignition experiment. Over a certain range of temperatures above the critical temperature they have found that kt_i/r^2 has a value of approximately 0·5 to 1. kt_i/r^2 only falls below 0·5 when $E(1/T_c - 1/T_i)$ exceeds 2, where T_c is the critical temperature and T_i is the actual ambient temperature in an experiment ($> T_c$). For values of E of order 30000 and $T_i \simeq T_c \simeq 500°$ K this means that $kt_i/r^2 > 0{\cdot}5$ if $T_i - T_c < 10°$ C say. kt_i/r^2 is greater than unity if $E(1/T_c - 1/T_i)$ is less than about 0·5, i.e. if $T_i - T_c < 2°$ C, and its value is then very sensitive to the precise value of $T_i - T_c$, tending to infinity as $T_i \to T_c$. A set of values of kt_i/r^2 observed under conditions where $T_i - T_c$ is between 2 and 10° C should therefore tend to lie about a mean value in the range 0·5 to 1. Small variations in the initial temperature T_i of the specimen will then be unimportant.

In Mitchell's experiments the highest non-ignition temperature and the lowest ignition temperatures were quoted. Since, in the majority of cases, the difference between them is between 2 and 10° C, kt_i/r^2 should be approximately constant in the range 0·5 to 1·0 for a wide range of r.

Mitchell's data for ignition, plotted in Fig. 5, do indeed follow a linear relation between t_i and r^2 very closely. For the wood fibre insulating board kt_i/r^2 equals 0·75 which lies in the above range. The discontinuity in the line for cane fibre insulating board in Fig. 5 occurs where a different construction was used by Mitchell for the smaller specimens and a different apparatus was employed.

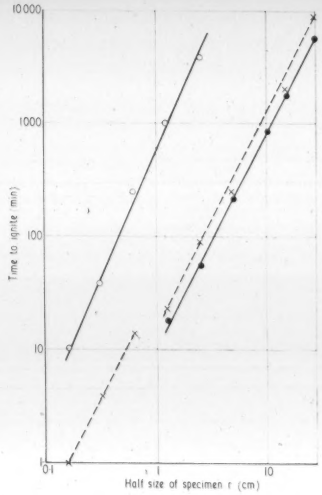


Fig. 5. Induction times at critical conditions (from Mitchell).
——○—— Cotton felted fibre insulating board.
- - - ×- - - Cane fibre insulating board.
——●—— Wood fibre insulating board.

Akita (1956) and Kinbara and Akita (1960) have reported experimental values of kt_i/r^2 of order unity for sawdust.

For Mitchell's results on cotton linters, kt_i/r^2 is about 40 times greater and, since this is systematic, it indicates a basic difference in behaviour. A difference such as this could arise if the reaction were autocatalytic or if reaction inhibitors were present. The induction time would then become long because of the time taken to form the active product material or to destroy the reaction inhibitors. From a practical point of view such a difference between the materials is important in the following respects: (i) it takes longer to judge whether the material is safe if the induction time is long, (ii) preheating would be expected to hasten ignition on subsequent heating for an autocatalytic reaction or if reaction inhibitors are initially present.

Discussion

The simple theory for thermal ignition assuming a single reaction with no loss of reactant gives relationships between the size of the specimen and the minimum temperature for ignition, and between the size of the specimen and the maximum temperature rise due to self-heating without ignition, which appear to be of the right form in so far as there is qualitative fit with Mitchell's results. But there is no quantitative agreement between the results of the two types of experiment.

Again there is disagreement, on the basis of the simple theory, between Mitchell's self-heating data and the results obtained in an adiabatic calorimeter by Gross and Robertson.

Theoretical considerations show that these discrepancies are greater than can be attributed to a low heat of reaction in one first order reaction. Experimental results show that, in fact, more than one exothermic reaction occurs in the self-heating of wood fibre insulating board and theoretical considerations show that in the first of these reactions the heat of reaction is too low to result in ignition.

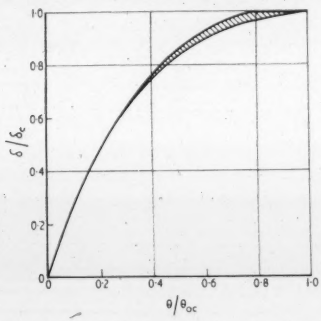


Fig. 6. The relationship between δ/δ_c and θ_0/θ_{0c} for all values of α.

It is suggested that Gross and Robertson's data agree with Mitchell's data for the ignition of wood fibre insulating board because in both cases the first, fast, reaction is completed at an early stage of the heating, so that in both experimental conditions only the second reaction is important. This may not be true for all materials. For example, with cotton linters which has an anomalous induction time, possibly due to more complex kinetics than the simple theory allows for, Gross and Robertson give an activation energy larger than that obtained from correlating Mitchell's ignition data.

Conclusions

The ignition data provided by Mitchell can be correlated by thermal ignition theory based on a single reaction obeying the Arrhenius law, though, of course, this does not imply that the reaction is as simple as this.

The rates of heat output for wood fibre insulating board calculated from the maximum self-heating temperature are approximately an order higher than those calculated from ignition data.

It is shown that the discrepancy cannot be accounted for solely by loss of reactant in the ignition experiments; but it can be partially explained by the presence of more than one exothermic reaction, for which some experimental evidence has been obtained.

The induction times are an independent source of data and the fact that there are anomalies between the results for the cotton felted linters and the two other materials tested by Mitchell strongly suggest that there is a basic difference between the reactions in this material and the others.

The main practical implications of this paper are that there are occasions when it is justifiable to use a simple model to extrapolate small-scale self-ignition data over a temperature and size range. Certain types of self-heating experiments may, however, give misleading predictions for self-ignition. It is advisable to perform both self-heating and self-ignition experiments on small specimens, and to find out if the two sets of data are consistent on the basis of the chosen theoretical model in order to assess the feasibility of extrapolation to full-scale.

Acknowledgment

The work described in this paper forms part of the programme of the Joint Fire Research Organization of the Department of Scientific and Industrial Research and Fire Offices' Committee; the paper is published by permission of the Director of Fire Research.

References

AKITA, K., 1956, *Bull. Fire Prev. Soc. Japan*, **5,** 43.
BERKOWITZ, *N.*, 1957, *Fuel*, **36,** 355.
BOWDEN, F. P., and YOFFE, A. D., 1952, *The initiation and growth of explosions in liquids and solids* (London: Cambridge University Press).
BREGER, I. A., and WHITEHEAD, W. L., 1951, *Fuel*, **30,** 247.
BUBEN, N. YA., 1945, *Zhur. Fiz. Khim*, **19,** 250.
BYRAM, G. M., FONS, W. L., SAUER, F. M., and ARNOLD, R. K., 1952, *Thermal Properties of Forest Fuels, U.S. Dept. of Agriculture Forest Service Division of Fire Research*, Oct.
CARSLAW, H. S., and JAEGER, J. C., 1947, *Conduction of heat in solids* (London: Oxford University Press).
CHAMBRÉ, P. L., 1952, *J. Chem. Phys.*, **20,** 1795.
CHANDRESEKAR, S., and WARES, G. W., 1949, *Astrophys. J.*, **109,** 551.
COOK, G. B., 1958, *Proc. Roy. Soc. A*, **256,** 154.
FRANK–KAMENETSKII, D. A., 1939, *Zhur. Fiz. Khim.*, **10,** 738.
—— 1955, *Diffusion and heat exchange in chemical kinetics* (Princeton; Princeton University Press).
GROSS, D., and ROBERTSON, A. F., 1958, *J. Res. Nat. Bur. Standards*, **61,** 413.
KINBARA, T., and AKITA, K., 1960, *Combustion and Flame*, **4,** 173.
MITCHELL, N. D., 1951, *Quart. Nat. Fire Prot. Assoc.*, **45,** 165.
RASKIN, W. H., and ROBERTSON, A. F., 1954, *Rev. Sci. Instrum.*, **25,** 541.
RICE, O. K., ALLEN, A. O., and CAMPBELL, H. C., 1935, *J. Amer. Chem. Soc.*, **57,** 2212.
RIDEAL, E. K., and ROBERTSON, A. J. B., 1948, *Proc. Roy. Soc. A.*, **195,** 135.
SEMENOFF, N. N., 1928, *Z. Phys. Chem.*, **45,** 571.
SMITH, K. A., 1959, *Tech. Ass. Pulp Pap. Ind.*, **42,** 869.
THOMAS, P. H., 1958, *Trans Farad. Soc.*, **54,** 60.
—— 1959, *Fire Research Abstracts and Reviews*, **1,** 143.
TODES, O. M., and MELENTIEV, P. V., 1939, *Zhur. Fiz. Khim.*, **13,** 1594.
TODES, O. M. *et al.*, 1933, *Zhur. Fiz. Khim.*, **4,** 71, 81.
—— 1939, *Zhur. Fiz. Khim.*, **13,** 868, 1594.
—— 1940, *Zhur. Fiz. Khim.*, **14,** 1026, 1447.
UBBELOHDE, A., and WOODWARD, P., 1948, *Phil. Trans A.*, **241,** 222.
ZINN, J., and MADER, C. L., 1960, *J. Appl. Phys*, **31,** 323.

Appendix 1

Relation between θ and δ for non-critical conditions in a sphere

The maximum theoretical value for the rise in temperature in self-heating is obtained by neglecting the loss of reactant.

The steady state equation for a sphere with no loss of reactant is, from Eqn (5*a*)

$$\frac{1}{z^2}\frac{d}{dz}\left(z^2\frac{d\theta}{dz}\right) = -\delta e^{\theta}. \tag{14}$$

Following Chambré (1952) we substitute in Eqn (14) $\eta = z(\delta \exp \theta_0)^{\frac{1}{2}}$ and $\psi = \theta_0 - \theta$. We then obtain

$$\frac{1}{\eta^2}\cdot\frac{p}{d\eta}\left(\eta^2\frac{d\psi}{d\eta}\right) = -e^{-\psi}. \tag{15}$$

At the centre, $x = 0$, $\theta = \theta_0$ and $d\theta/dz = 0$.

Therefore at $\eta = 0$, $\psi = 0$, $d\psi/d\eta = 0$ (16)

and at the boundary $z = 1$, $\alpha\theta = -d\theta/dz$. Therefore at $\eta = (\delta \exp \theta_0)^{\frac{1}{2}} = \eta_s$

we have $$\alpha(\theta_0 - \psi_s) = \eta_s(d\psi/d\eta)_s. \tag{17}$$

Tables of ψ as a function of η for Eqns (15) and (16) are given by Chandresekar and Wares (1949). The values of ψ_s and α satisfying Eqn (17) for given δ and θ_0 can then be obtained.

Hence δ can be obtained as a function of θ_0 and α. The results are shown in Fig. 2.

Although there is a large variation of δ with α for a given θ_0, a curve of δ/δ_c against θ_0/θ_{0c} is practically independent of α (Fig. 6).

Small values of θ_0

For small values of θ_0 we may put θ equal to θ_0 on the right-hand side of Eqn (14) which can then be integrated to give

$$\frac{d\theta}{dz} = -\frac{\delta z \exp \theta_0}{3}$$

and $$\theta = A - \frac{\delta z^2 \exp \theta_0}{6}$$

so that from (2*c*)

$$\alpha(A - \tfrac{1}{6}\delta \exp \theta_0) = \tfrac{1}{6}\delta \exp \theta_0$$

and, at the centre $z = 0$,

$$\theta_0 = A = \frac{\delta}{6\alpha}(2 + \alpha) \exp \theta_0.$$

This relationship is shown in Fig. 2 and it is seen to be satisfactory for $\theta_0 < 0{\cdot}4$.

Putting $\exp \theta$ equal to unity on the right-hand side of Eqn (14) we find

$$\theta_0 = \frac{\delta}{6\alpha}(2 + \alpha)$$

which would appear as straight lines on Fig. 2 tangential at the origin to the correct solution. If we use this equation to plot Mitchell's self-heating data as

$$\ln\left(\frac{\alpha}{2+\alpha}\cdot\frac{T_0 - T_A}{r^2}\right) \text{ against } \frac{1000}{T_A} \text{ instead of } \frac{1000}{T_0}$$

we lessen the curvature in each set of data, but the common asymptote drawn in the region where $T_0 - T_A \to 0$ would be the same.

At large values of θ_0 the former approximation is the better. The curvature in each set of data corresponds to a real difference in the rate of heat generation in specimens of different size at the same temperature.

Appendix 2

Discussion of temperature–time curve in self-heating

We assume θ to be small compared with unity so that the rate of generation of heat is solely a function of the time under the nearly isothermal conditions of self-heating.

The self-heating equation may therefore be written as

$$\nabla^2\theta = \frac{d\theta}{d\tau} - \phi(\tau) \tag{18}$$

where $\phi(\tau)$ is the rate of generation of heat

and where $$\int_0^\infty \phi(\tau)d\tau = B. \tag{19}$$

For symmetrical heating and boundary conditions at each surface of the form given by equation (2*c*), the maximum temperature, i.e. the central temperature in a cube, denoted by θ_0, can be readily calculated by the methods described by Carslaw and Jaeger (1947, §132–4). From this solution we obtain

$$\int_0^\infty \theta_0 d\tau = MB$$

i.e. $$\int_0^\infty (T_0 - T_A)dt = \frac{MQ\lambda_0 r^2}{K}.$$

M is given by

$$M = 8\sum_{m=1}\sum_{n=1}\sum_{p=1}\frac{1}{(\beta_m^2 + \beta_n^2 + \beta_p^2)} \Pi_{i=m,n,p}\left\{\frac{\beta_i \sin \beta_i\{\cos \beta_i + (\alpha/\beta_i)\sin \beta_i\}^2}{\alpha + \alpha^2 + \beta_i^2}\right\} \tag{20}$$

where β is a solution of $\beta \tan \beta - \alpha = 0$.

$\Pi_{m,n,p}$ denotes the product over the terms in m, n, p. This result is independent of the form of $\phi(\tau)$.

The second reaction may be approximated by a first order reaction for which we have

$$\phi(\tau) \simeq a\delta \exp\left(-\frac{a\delta\tau}{B}\right) \tag{21}$$

where a is nearly equal to unity, and the solution takes the form

$$\theta_0 = \sum_m\sum_n\sum_p b_{m,n,p}\left[\exp\left(-\frac{a\delta\tau}{B}\right) - \exp\{-(\beta_m^2 + \beta_n^2 + \beta_p^2)\tau\}\right] \tag{22}$$

where the coefficients $b_{m,n,p}$ depend on the roots $\beta_{m,n,p}$, α and $a\delta/B$.

Now all the roots β are greater than 1 if $\alpha > 2$. If also $a\delta/B \ll 1$ and $\tau > 0{\cdot}8$ only the first of the two exponential terms is significant, i.e. $\theta_0 \propto \exp(-a\delta\tau/B)$. Similarly for other orders of reaction, provided ϕ changes slowly with τ compared with $\exp(-3\beta_1^2\tau)$ the solution takes the asymptotic form $\theta_0 \propto \phi(\tau)$, i.e. there is no thermal delay in the response of the temperature of the material to a changing rate of heat output when τ is greater than a certain minimum value.

Some visco-elastic properties of nylon

by F. L. WARBURTON, J. F. P. JAMES* and (Miss) S. HOTHER-LUSHINGTON, Wool Industries Research Association, Headingley, Leeds, Yorks.

MS. received 24th August, in revised form 5th December 1960

Abstract

The dynamic modulus of nylon fibres has been measured over a wide frequency range at different temperatures and humidities, and also the life under load. The measured life times were of the same order as the periods used in the dynamic experiments. The results of both sets of experiments can be explained on the assumption that increase in temperature only affects the internal friction of nylon, whereas the absorption of water has also a significant although smaller effect on the elastic properties.

1. Introduction

In a previous paper (Warburton 1959) one of the authors gave an account of an investigation into the dynamic mechanical properties of wool at very low frequencies. In this paper it was shown that the behaviour could best be characterized by a visco-elastic assembly in series with a pure spring; both the elastic and viscous elements of the required assembly were affected in a uniform manner by absorption of water, whereas the series spring was unaffected. The effect of temperature, on the other hand, only required an alteration in the viscous elements of the model for its description. In the present paper the results of a similar investigation into the dynamic properties of nylon are reported, and also those of a simultaneous investigation into the effects of humidity and temperature on the life of a nylon fibre under load.

As in the previous paper it has been found possible to fit all the measurements of the dynamic compliance to a single curve, by applying appropriate reduction factors to the time and compliance scales. If it is assumed that the distribution of breaking times for a large number of fibres under any given conditions merely reflects a quite reasonable distribution of minimum effective cross sections, it is also possible to fit all the breaking-time data to a single curve by applying similar reductions to the load and time scales. The range of breaking times studied was very similar to that of the periods of vibration used in the dynamic experiments; it is therefore possible to compare the shifts required in the two investigations.

2. Experimental

The dynamic mechanical properties at low frequencies were studied by making simultaneous measurements of force and elongation during longitudinal vibrations. The apparatus used was almost identical with that previously described (Warburton 1959). Minor modifications were made to extend the range that could be covered on the high-frequency side. The nylon fibre used was cut from a bobbin of 12 denier monofilament yarn supplied by British Nylon Spinners. A fairly coarse fibre was required so that the stiffness of the spring used for measuring the force was sufficiently great to keep the resonance frequency well above the maximum experimental value. The experimental procedure was identical with that previously described and need not be discussed further.

The life of nylon fibres under load was studied by suspending about fifty fibres in a controlled atmosphere and determining the time at which each fibre broke. The controlled atmosphere was obtained by enclosing the whole apparatus in a box (Fig. 1), in which the temperature was thermostatically controlled and the humidity controlled by trays containing salt solutions. A multi-fan system was used to provide adequate internal circulation, so that uniform conditions were attained throughout the box and close control was possible. The variation of temperature throughout the box was within $\pm 0{\cdot}2$ deg C of the mean temperature, which itself varied by $\pm 0{\cdot}2$ deg C, and the humidity was controlled to $\pm 2\%$ r.h. The fans were so arranged that the air movement in the vicinity of the fibres, while sufficient for maintaining uniform conditions, was too small to disturb the fibres when under load. The fibres were suspended from hooks on the removable carrier A (Fig. 1) by special clamps made from safety pins and a Paxoline pad. The design of

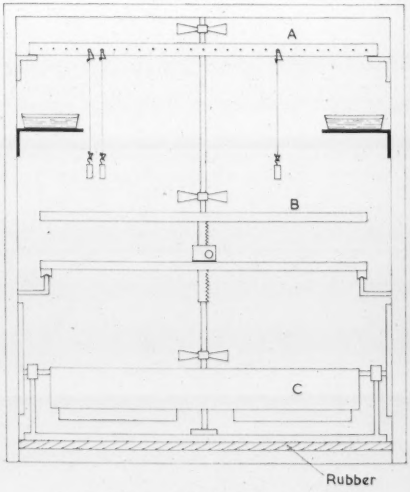


Fig. 1. Diagram of breaking time apparatus.

* Now at Commonwealth Scientific and Industrial Research Organization, Wool Research Laboratories, Division of Textile Physics, Ryde, New South Wales, Australia.

these clamps is shown in Fig. 2, an important feature being the wrapping of the fibre round the wire; this produces a capstan action by which part of the load is carried by static

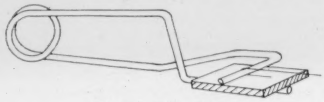


Fig. 2. Fibre clamp.

friction between the fibre and wire, thereby reducing the chance of a break occurring at the nip. Similar clamps were fitted at the lower end at a distance of about 6 cm, a special jig being used to facilitate attachment at the correct distance.

At the start of each experiment the appropriate number of weights was placed on the tray B (Fig. 1); this was raised by a rack and pinion so that the weights could be hooked to the lower clamps without applying any tension to the fibres. When all the weights were in position the box was closed and the fibres left to condition for at least four hours, usually overnight. After the fibres had been conditioned, the tray supporting the weights was gently lowered and pushed back clear of the suspended weights. It was essential to avoid any jarring during these operations as any vertical oscillation of the weights would add to the stress on the fibres. To assist in preventing such oscillations much of the apparatus was mounted on rubber. When a fibre broke, the falling weight struck the pivoted plate C, which then tilted and completed an electric circuit, thereby deflecting a pencil in contact with a moving strip of paper. The paper was driven by a synchronous clock motor which was started at the same time as the platform was lowered, so that the distance traversed by the paper before each mark was made was a measure of the breaking time of the corresponding fibre. Only the shorter breaking times were recorded in this way, although the first few overnight breaking times were recorded with the paper running at a much slower speed. Other breaking times were recorded either with a stop watch or an ordinary watch, according to the time that had elapsed since the start of the experiment. The later overnight breaking times were determined with sufficient accuracy by assuming that they occurred at evenly spaced intervals. Thus even on the third night (after about 50 hours) the maximum error in log breaking time caused by assuming a single fibre broke at 1 a.m. would be $\pm 0{\cdot}06$, the error becoming progressively smaller as the time elapsed increased; with the smaller loads and low humidities the experiment was stopped before the stronger fibres had broken. Even so lifetimes of several weeks were recorded.

The nylon fibres were drawn from a sample of 3 denier nylon tow having a mean diameter of 20 μ, and the experiments were made using loads of 11·93, 9·50 and 7·64 g at 25 and 40° c and various humidities. All the weights corresponding to any one load were adjusted, by filing or otherwise, in such a manner that the variability was small compared with that of the fibre diameter, the standard deviations being 56, 8 and less than 4 mg respectively. Except for the 11·93 g weights these are negligible, and even for these it corresponds to a variance of less than $\frac{1}{2}\%$.

3. Experimental results

The variation of the real part of the dynamic compliance with frequency for different regains is shown in Fig. 3. The

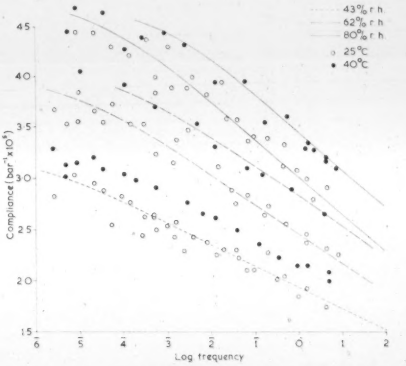


Fig. 3. Relation between real part of dynamic compliance and frequency.

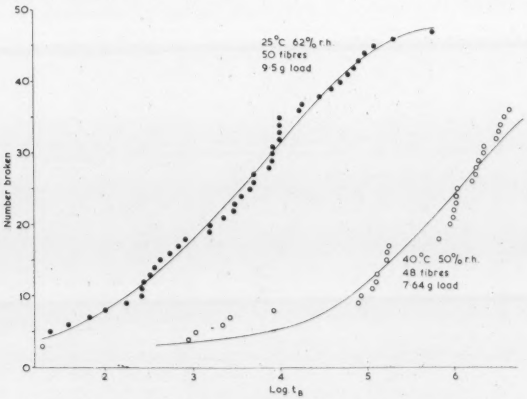


Fig. 4. Relation between number of fibres broken and time.

curves are very similar to those for wool (Warburton 1959) but extend over a greater frequency range. They show that the low-frequency limit at least is humidity dependent, and that master curves can only be obtained by applying a reduction factor to the compliance scale as well as to the frequency scale. Further discussion of these data is deferred to the next section.

The data from typical experiments on the life under load are shown in Fig. 4. In this figure the number of fibres broken is plotted as ordinate against the logarithm of the time in seconds. It will be noted that the breaking times cover a very large range, there being a factor of over 10^5 between the longest and shortest breaking times in any one experiment. It will also be noticed that the points do not lie on a smooth curve but fall in a series of steps (almost straight lines). Van der Vegt (Stuart 1956) obtained similar curves for the lifetime of Nylon 66 fibres subjected to repeated torsional deformations. The dependence of the breaking time on load is shown in Fig. 5, in which the load is plotted against the logarithm of the half-life, i.e. the time for half the fibres to break. This figure shows the large effect of even a small change in load; thus a change of less than 7% in the load alters the half-life by a factor of 10 and a change of $1\frac{1}{4}$% would make a difference of 0·2 in the logarithm of the half-life. The variability of weights ($\pm 2\,\sigma$) is less than ± 1% even for the heavier weights, so that this factor only makes a small contribution to the scatter of the results. It will, however, affect the individual breaking times under the 11·93 g load, but the effect will be small for the 9·50 g load and negligible for the 7·64 g load.

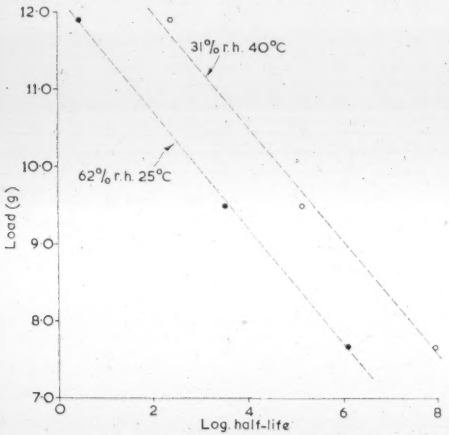


Fig. 5. Relation between half-life and load.

4. Discussion of the results

In determining the half-life under load (or any other single parameter that could be used to summarize the results of each test) very little use is made of the data relating to the scatter of breaking times. If a correlation is to be made with the modulus data, some means of applying a reduction factor to both the load and time axes for different regains is required. This can only be done if analysis is made of the distribution of breaking times in each experiment. Another reason for so doing is the fact that the similarity of the curves in Fig. 4 to those obtained by van der Vegt (Stuart 1956) suggests that the way in which the points are scattered round the smooth curve is not an artefact but has some real significance, even if only statistical.

A possible reason for the large variation in breaking times is a random distribution of weak places. If this is expressed as a variation in effective areas under the applied load, the variance required in the minimum value is not large and could easily be produced by a random distribution of sub-microscopic flaws. (The small amount of delustering material found in all commercial nylon would contribute to such flaws.) These need not be on the surface of the fibre; in fact, microscopic examination of the broken ends shows the cup and ball fracture characteristic of the presence of interior flaws. On the above hypothesis, the scatter of the points with respect to a smooth curve might be explained by bunching taking place when a finite number of individuals are drawn at random from an infinite sample having a true Gaussian distribution.

The variance in effective area at the point of break, and hence in equivalent diameter, required to give the observed distribution can be calculated from curves of the types shown in Fig. 4 and the relation between load (and hence stress) and half-life as shown in Fig. 5. Thus curve A in Fig. 4 relates to a test in which fifty fibres were used; 68% of these would be expected to have diameters differing from the mean within $\pm\sigma$, the standard deviation, i.e. those from the 8th to the 42nd fibre broken. The log breaking times for these two fibres were 2·0 and 4·88 respectively, and from Fig. 5 one finds, assuming for simplicity of calculation a linear relation between log stress and log breaking time,

$$\log (W/d^2) = \text{const.} - 0{\cdot}03 \log t \text{ (approximately).}$$

Hence

$$\log \{W/d_8^2\} - \log \{W/d_{42}^2\} = 0{\cdot}0864.$$

Therefore if c is the coefficient of variation σ/d_{mean}

$$\log (1 + c)/(1 - c) = 0{\cdot}0432, \quad \text{whence } c = 0{\cdot}05.$$

Several such calculations were made and gave a mean value of 5% for the variance within the limits of the accuracy obtainable, this being limited by the accuracy of the relation between load and median breaking time. Using this provisional value of the variance, an integrated Gaussian curve was plotted from the tables given in the Handbook of Chemistry and Physics (Cleveland, Ohio: Chemical Rubber Publishing Co.) as follows: the argument t is equal to $d - 20$ and the area of the error curve (as given in these tables) multiplied by the total number of fibres is equal to the number of fibres having diameters between 20 μ and d. If it is then assumed that the fibres break in order of minimum load bearing area, the data of Fig. 4 can be converted to a relation between W/d^2 and log breaking time. This has been done in Fig. 6; the fact that the points for different loads under the same conditions lie as nearly as possible on a single curve shows that the value adopted for the variance is as near as can be obtained from the experimental data; i.e. the breaking times for the same calculated stress with different loads agree to within the general scatter.

In order to further test the hypothesis outlined above—in particular, the attribution of the scatter of the points round a smooth curve to a bunching effect—theoretical life under load data for forty fibres have been obtained. In so doing it has been assumed that the breaking time is solely determined by the maximum stress in each fibre, and that the

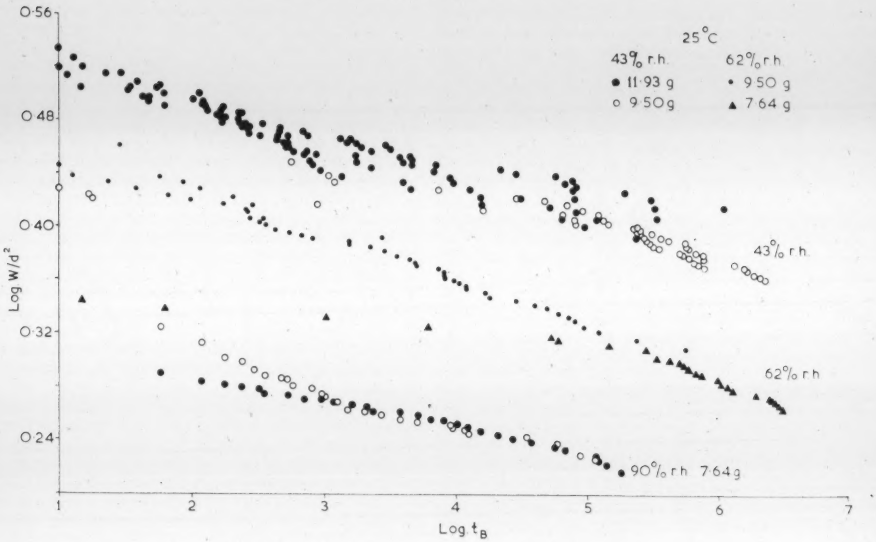


Fig. 6. Relation between maximum stress and breaking time.

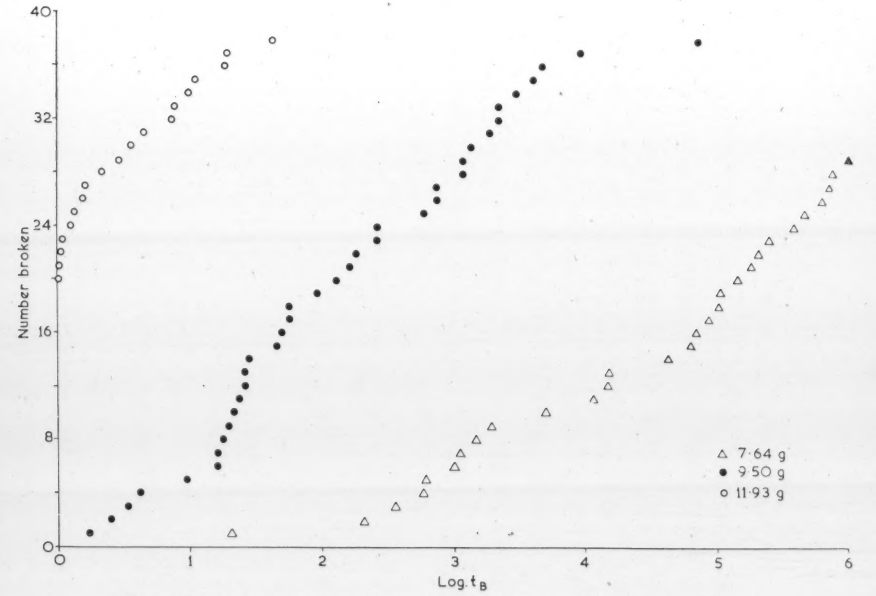


Fig. 7. Relation between number of fibres broken and time from theoretical fibre draws.

effective minimum diameters of the forty fibres were those to be expected from a random sample drawn from an infinite population having a Gaussian distribution of this parameter. For convenience the numerical values used were similar to those used above, i.e. $c = 0{\cdot}05$ and mean diameter $= 20\ \mu$, hence the standard deviation is $1\ \mu$. The deviations from the mean for each of the forty fibres can then be read directly from the tables of random samples from a normal distribution, such as those produced by P. R. Ray (1934). (The method used in producing these tables was given by Karl Pearson in his introduction to Tippet's (1927) *Random Sampling Numbers*.) The breaking times for each fibre under a given load were then calculated from the relation

$$0{\cdot}03 \log t = 0{\cdot}433 - \log \text{stress}.$$

The resultant curves are shown in Fig. 7 which refers to three separate draws, different applied loads being assumed for each draw. These curves exhibit all the main features of the experimental curves as shown in Fig. 4, and thus show that the departure from a smooth curve produced by the random selection of a small number of items is quite capable of producing the observed scatter of the individual points, although one could not assert that the curves in Fig. 4 arose from a single Gaussian distribution.

The curves of Fig. 7 have been converted back to a relation between log stress and breaking time using the same procedure as used in constructing Fig. 6, i.e. assuming that the fibre diameters were uniformly distributed according to the Gaussian relation; the results are shown in Fig. 8. The breaking times from this cause, since there are two factors known to operate, which would contribute to the variation from fibre to fibre of the minimum load bearing area. There is, in the first place, a variance of 5% in fibre diameter as measured by the standard microscope technique which will produce a contribution much less than 5%, and secondly, there is a random distribution of delustering particles which is probably the major factor. There may be other factors.

The above analysis has therefore established in the first place that known sources of variability can account for the extreme variation from fibre to fibre in life under load, without assigning unreasonable values to their magnitude. It is therefore not necessary to invent special theories of fracture to explain this variability; i.e. on the basis of these experiments, all that is required of a theory of fracture is that it should account for the large variation in breaking time with stress and for the effects of absorbed moisture and temperature on this property. More important, it does show that whatever the ultimate variation between fibres, which gives rise to the variability in breaking times, the scatter of the experimental points from a smooth curve almost certainly arises mainly from statistical causes rather than from the mechanism of fracture.

The procedure adopted does enable one to use all the data to obtain a relation between maximum applied stress and breaking time with a reasonable degree of confidence, whatever the ultimate source of the variability. Such a relation is required before one can discuss the effects of temperature and absorbed moisture on the breaking time or compare

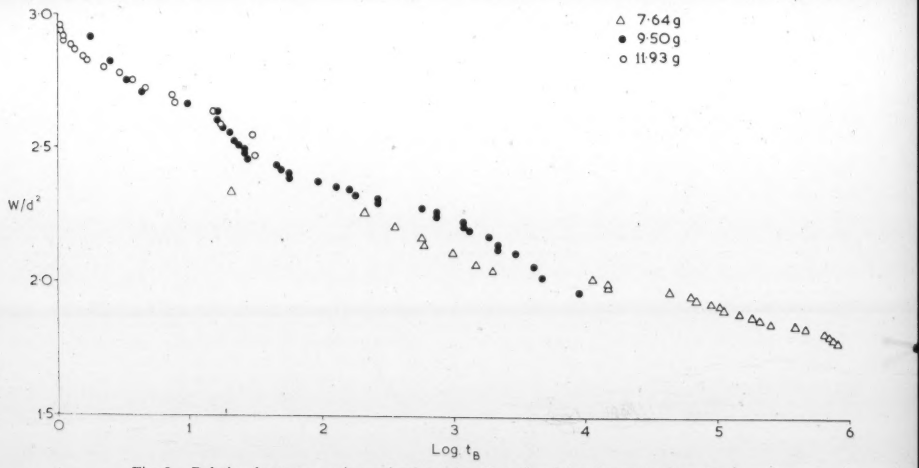


Fig. 8. Relation between maximum load and breaking time from theoretical fibre draws.

points for the different loads do not show a much better fit to a single curve than do those in Fig. 6, the differences between the two arising possibly from a slightly non-Gaussian distribution. It follows that if the variation in experimental breaking times is attributed solely to a variation in minimum load bearing area, the variance of the best Gaussian distribution of diameters (used as a measure of this parameter) has a value of 5% to within the limits of experimental accuracy. There must be a contribution to the variability of them with the effects of these factors on the dynamic modulus. One way of discussing these effects is to reduce all the data to a single curve, by sliding the results for different ambient conditions along the log stress and log breaking time axes; the two shifts can then be interpreted as arising from changes in the strength of the structural elements and in the internal fluid friction. Unfortunately the curves in Fig. 7 are so nearly straight and parallel that an almost infinite combination of shifts can be used. The most helpful tentative procedure is

to assume that the changes in internal friction, as found from measurements of the dynamic modulus, hold for extensions up to fracture, and then to calculate the change in strength required to produce complete coincidence. Further analysis must therefore be deferred until the dynamic modulus data have been examined.

Reverting now to Fig. 3, relating the dynamic compliance to frequency, it is quite clear from even a casual inspection of the curves that it is impossible to construct a master curve for different regains simply by sliding the curves along the log frequency axis. The same procedure must therefore be adopted as was used for wool (Warburton 1959), and a reduction factor applied to the compliance scale as well as to the frequency scale. There are insufficient data from these experiments to decide whether this should be applied to the whole of the compliance or merely to the amount by which the actual compliance J_ω exceeds J_∞, the high-frequency limit to which the compliance would be expected to tend. Recourse must therefore be had to other data of which those obtained by Quistwater and Dunell (1958, 1959) are suitable for this purpose. These workers, using an entirely different technique, determined the real and imaginary parts of the dynamic modulus of nylon 66 over a decade of frequency overlapping the high-frequency end of the present investigation. Their data, however, extend to lower humidities and regains than those reported here, and a constant value for the modulus was obtained for relative humidities below about 20% at a temperature of 9° C. This limiting value is about $7{\cdot}5 \times 10^4$ bar, and the loss modulus was about 10^4 bar, so that the limiting value of the compliance would be about $1{\cdot}3 \times 10^{-5}$ bar^{-1}. Using this value the data for 25° C have been reduced to a single curve (Fig. 9) by applying appropriate

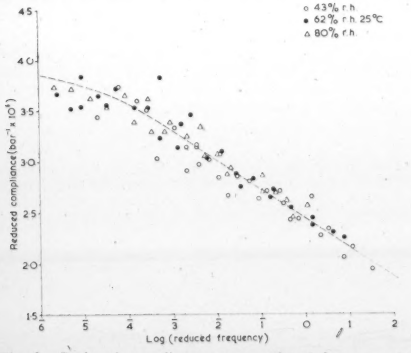


Fig. 9. Reduced compliance versus reduced frequency for 25° C.

reduction factors to the time scale and the excess compliance above $1{\cdot}3 \times 10^{-5}$ bar^{-1}. In doing this the curve for 3·85% regain (62% r.h.) at 25° C has been taken as standard, the reduction factors for 2·6% regain and 5·6% regain at 25° C being +0·9 and −0·6 respectively in log frequency and +0·165 and −0·11 in the logarithm of the excess compliance above $1{\cdot}3 \times 10^{-5}$ bar^{-1}. In their second paper, Quistwater and Dunell (1959) constructed master curves for 35° C and 60° C by simply sliding the points for each humidity along the log frequency axis, a procedure which is in apparent contradiction to the results given above. However, when the experiments only cover a short frequency range, it is possible to construct equally satisfactory master curves by applying a variety of combinations of reduction factors to the compliance (or modulus) and frequency. Quistwater and Dunell naturally chose the simplest procedure in the absence of evidence to the contrary.

The curves for 40° C, in Fig. 3, can be brought into coincidence with the corresponding curves for 25° C by displacing them along the log frequency axis by 1·2 towards the lower frequencies.

Only the real part of the dynamic compliance has been considered above. No satisfactory analysis of the loss compliance has been possible. This is because the amplitude of the deformation cycles was large enough for the shape of the hysteresis loop (Fig. 10) to be affected by non-linearity

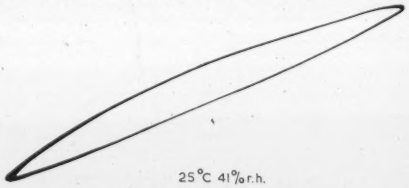


Fig. 10. Experimental hysteresis loop.

of the elastic elements in the nylon. If the behaviour could be characterized by linear visco-elastic elements over the range studied, these should be ellipses no matter how complicated the model used, whereas it will be seen that they depart appreciably from this shape. Although analysis in terms of a linear model probably leads to correct values for the mean slope of the load elongation for the elastic elements of the model over the experimental range, the interpretation of the width of the loop becomes more complicated and requires some expression for the departure from non-linearity, although sufficient accuracy can be obtained for the small correction required to obtain the real part of the compliance. Solution of the differential equation for the behaviour of an assembly of non-linear visco-elastic elements is extremely difficult, but it is possible to determine the behaviour of a single non-linear Voigt element of a suitable type; the curve in Fig. 11 shows the type of hysteresis loop that can be obtained from the Burte–Halsey model (1947, see also Peters and Speakman 1948). The distortion of the loop from an ellipse can be reversed according to whether one considers deformation in a region before or after the point of inflection in the spring line of the model. There is no reason to doubt that the same would apply to a model using any other spring line showing a point of inflection. Examination of the loops obtained from nylon show that they all correspond to a spring line of the type found after the point of inflection of the spring line, i.e. spring line concave towards the load axis. The hysteresis loops obtained for wool (Warburton 1959) departed less from the elliptical form and distortions in both directions were found. One can conclude therefore that the experiments on wool were made in a region close to the point of inflection of the spring line. Another type of spring line which shows the necessary characteristics is that given by the statistical theory of elasticity for the non-Gaussian network (Stuart 1956, p. 325). On either model the comparison of the results for nylon and wool would suggest that nylon 66 fibres are much more highly oriented than those of wool.

Another consequence of the non-linear behaviour of the nylon is that the calculated compliance will depend to a

certain extent on the mean extension. Although an endeavour was made to keep this constant for all measurements, it could only be kept approximately so and this

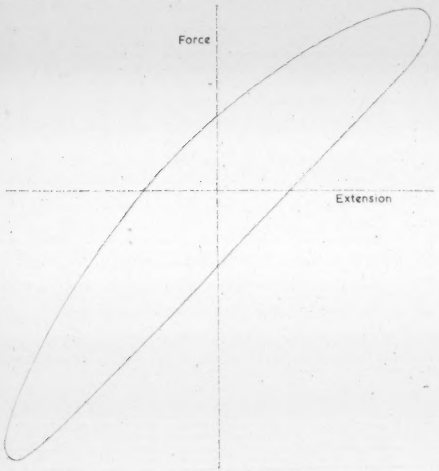


Fig. 11. Hysteresis loop for Burte–Halsey element.

factor will contribute to the scatter of the points in Figs 3 and 9.

The reduction factors which have been already used to construct master curves for the dynamic modulus data can also be used as a guide in constructing similar curves for the breaking time data. In so doing it is assumed that the reduction factors for internal viscosity, which have been used for the dynamic modulus data, apply also to the breaking time experiments, despite the fact that one factor applies to small extensions and the other to all extensions up to break. If this is done any remaining difference between the data for different ambient conditions is attributed to a change in the strength of the structural elements. With the breaking time data the changes in breaking stress required to produce coincidence are likely to be small, so that the shifts to be applied to the log time scale will be determined almost entirely by changes in viscosity. This is in contrast to the dynamic modulus experiments, where the changes in compliance were sufficiently large to have an appreciable influence on the change required in the log frequency scale. The factors to be applied to the log time axis in the breaking time experiments should therefore be compared with the sum of the factors used in the modulus experiments.

The reduced curves of log stress and log breaking time shown in Figs 12 and 13 have been obtained by the use of reduction factors which correlate well with each other and also with those used in constructing the curve in Fig. 9. In all three curves the data have been reduced to 62% r.h. (3·85% regain). Further, the curves in Figs 12 and 13 differ only in that the curve in Fig. 13 is displaced a distance 1·2 towards lower log breaking times than that in Fig. 12, thus correlating the effect of temperature on the two phenomena, i.e. dynamic modulus and life under load. The log reduction factors used for the two axes are shown in Fig. 14 (*a* and *b*); Fig. 14(*a*) also shows the sum of the log factors used for constructing Fig. 9 (the sign being changed since log frequency is negative log time). It will be seen that all the values fit a single curve within the limits of experimental accuracy.

One can make, therefore, the generalization that the effect of temperature on both the dynamic modulus and life under load of nylon fibres arises from a reduction in internal fluid friction of such a magnitude that an increase of 15 deg C in temperature (from 25–40° C) reduces the friction in the ratio 1 : 15. The effect of moisture is more complicated in that, in addition to a decrease in fluid friction, additions of further moisture reduces both the elastic properties and strength of the structural elements. The curve in Fig. 14(*a*)

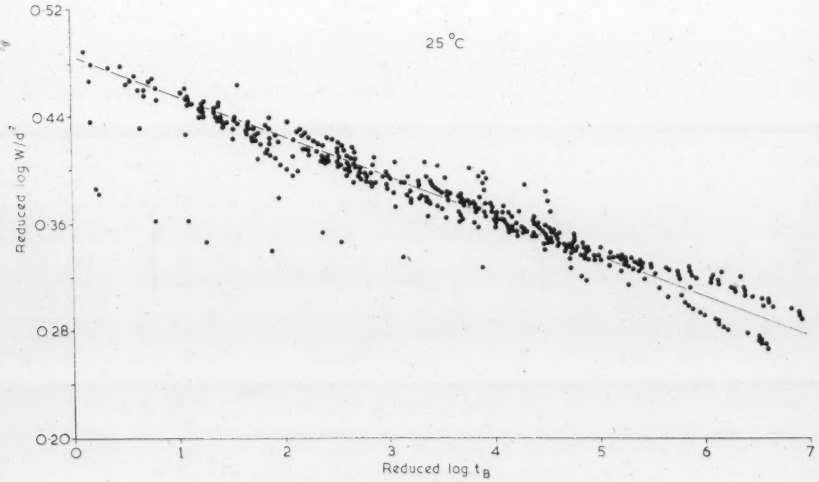


Fig. 12. Reduced stress versus reduced breaking time for 25° C.

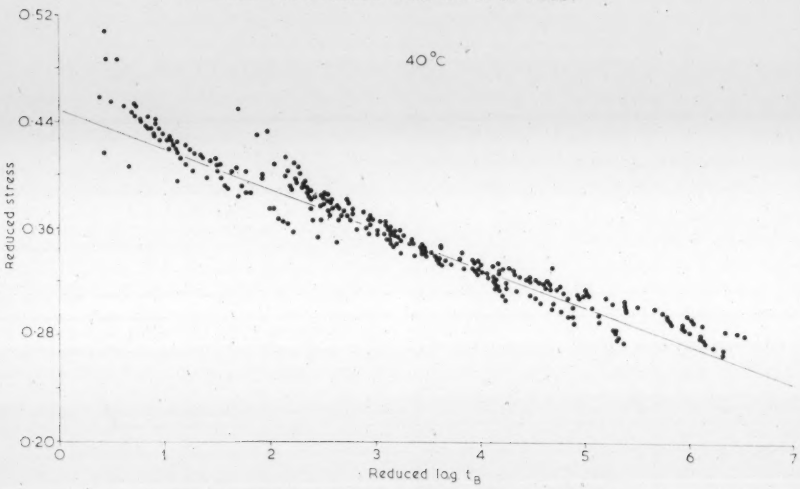


Fig. 13. Reduced stress versus reduced breaking time for 40° C.

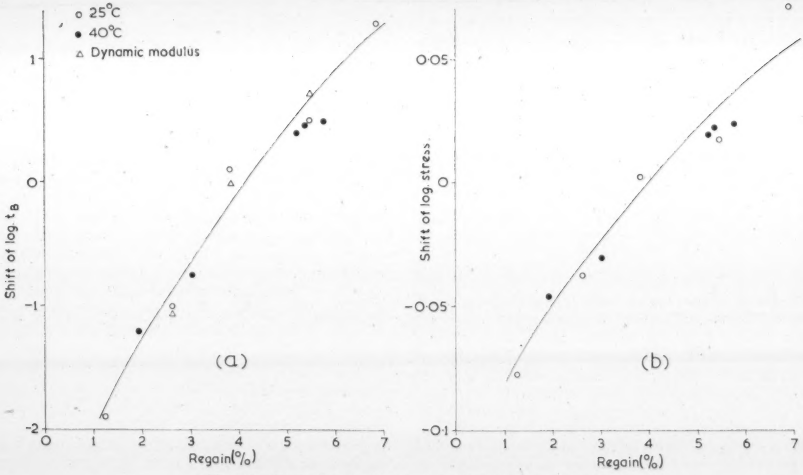


Fig. 14. Reduction factors for breaking time experiments.

may be taken as giving a relation between log (reciprocal of viscosity) and regain.

Acknowledgments

The authors are indebted to Dr. D. G. Medley of the Statistics Department for the suggestion that the step-like shape of the breaking time curves could have a statistical origin and for assistance in the statistical calculations. They are also grateful to British Nylon Spinners for supplying the nylon sample.

References

BURTE, H., and HALSEY, G., 1947, *Textile Research J.*, **17**, 465.

PETERS, L., and SPEAKMAN, J. B., 1948, *J. Textile Inst.*, **39**, P253.

QUISTWATER, J. M. R., and DUNELL, B. A., 1958, *J. Polymer Sci.*, **28**, 309.

—— 1959, *J. Appl. Polymer Sci.*, **1**, 267.

RAY, P. R., 1934, *Sankhyā: Indian J. Statistics*, **1**, 289, P. C. Mahalanobis *et al.*

STUART, H. A., 1956, *Die Physik der Hochpolymeren*, Vol. 4, 174 (Berlin: Springer).

TIPPETT, L. C., 1927, *Tracts for Computers* No. XV, p. iv (Cambridge: The University Press).

WARBURTON, F. L., 1959, *J. Textile Inst.*, **50**, T1.

Photographic technique for the determination of metal cutting temperatures

by G. BOOTHROYD, Royal Technical College, Salford, Lancs.

MS. received 2nd January 1961

Abstract

An experimental technique has been developed for the measurement of the temperature distribution in the deformation zone and in the chip and tool during orthogonal cutting. The term orthogonal cutting is used for the special case where the cutting edge of the tool is perpendicular to the direction of relative motion of the workpiece and tool.

The method involves photographing the chip, tool and workpiece in a plane perpendicular to the cutting edge, using an infra-red sensitive photographic plate. The optical density of the plate is determined over the relevant field by means of a microdensitometer. A heated tapered strip, on which the temperature distribution is measured by means of a series of thermocouples, is photographed simultaneously with the tool and workpiece. This enables the optical density of the plate to be calibrated in terms of temperature, and the temperature distribution over the chip, tool and workpiece can hence be determined.

The effect of surface emissivity and of other factors on the accuracy of the technique are considered, and the results of an experiment using the photographic technique are presented and discussed in relation to previous theoretical and experimental work.

1. Introduction

One of the important factors in metal cutting is tool wear and this is thought to be closely related to the temperatures existing at the wearing surfaces of the tool. Friction at the tool surfaces is also likely to be affected by these temperatures. For these reasons a number of investigations have been carried out in the past into the temperatures developed during metal cutting.

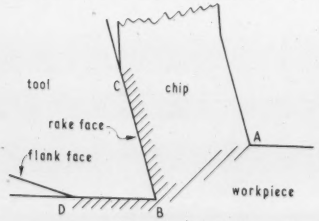


Fig. 1. Orthogonal cutting.
/// heat sources.

In the orthogonal cutting of a continuous chip (Fig. 1) the work of cutting appears as three main heat sources; heat generated over the region AB, known as the shear zone, or shear plane, due to the deformation of the work material; frictional heat generated in the region of the chip–tool interface CB due to the chip sliding over the tool rake face; frictional heat at the work–tool interface DB. With a perfectly sharp cutting tool this last heat source may not be present. This paper describes a photographic technique which has been developed for the measurement of temperatures generated in the tool, chip and workpiece.

2. Previous work

In previous experimental and theoretical work the temperatures in the shear zone, chip and tool have usually been considered separately.

Among the analytical studies of shear zone temperatures, Weiner (1955), by assuming that heat conduction in the workpiece in the direction of motion could be neglected, solved the heat transfer equation for the temperature distribution in the workpiece when the shear zone constituted a plane heat source of uniform strength. Rapier (1954) used a relaxation technique to obtain the temperatures in the workpiece with the same assumptions regarding the shear zone. The results of both workers agreed regarding the average value of the shear zone temperature.

Among the experimental studies of the temperatures in the workpiece, Nakayama (1956) used a thermocouple technique to determine the proportion of the shear zone heat conducted into the workpiece, and thus obtained an indirect measure of the average shear zone temperature. His results agreed well with the predictions of Weiner and Rapier. Reichenbach (1958) embedded thermocouples in the workpiece and recorded the temperature transients as they passed through the shear zone into the chip. From these results he was able to plot the temperature distribution along the shear zone. Hollander, however, in the discussion of this paper, pointed out that it was likely that the thermocouples would not maintain their accuracy owing to the severe deformation to which they would themselves be subjected. Hollander (1959) and Hollander and Englund (1957) used first thermocouples and later a microradiation pyrometer to determine the temperature distribution in the workpiece, but unfortunately made no measurements in the region of the shear zone.

In studying analytically the temperature distribution in the chip, Rapier (1954) assumed that heat conduction in the direction of motion could be neglected and that the heat source in the region of the chip–tool interface was a plane heat source of uniform strength along its length. His results showed that the maximum temperatures occurred on the rake face of the tool some distance from the cutting edge. This is the region where 'crater' wear of the tool takes place; the results, therefore, agree qualitatively with the predictions of

Trent (1952) based on a consideration of the mechanism of tool wear. Reichenbach's thermocouple technique gave temperatures in the chip in qualitative agreement, though much lower in magnitude, than those predicted by Rapier.

Analytical work of Chao and Trigger on temperatures in the cutting tool showed maximum temperatures on the rake face some distance from the cutting edge, and later work (Chao and Trigger 1958) included the effect of a frictional heat source at the flank face.

In the work-tool thermocouple technique (Herbert 1926) the e.m.f. generated at the junction of the tool and workpiece and chip indicates only an average temperature over the areas of contact. The temperature readings obtained by this method would be expected to be substantially lower than the maximum temperatures predicted by analytical methods.

3. Experiment

The principle of the technique was that of obtaining an infra-red photograph of the side surfaces of the workpiece, chip and tool during orthogonal cutting and simultaneously a photograph of a heated calibration strip, the temperature distribution of which was obtained by independent means. The photograph could then be calibrated for temperature distribution by means of the picture of the furnace.

The apparatus employed in the technique is shown schematically in Fig. 2. The cutting operation was orthogonal,

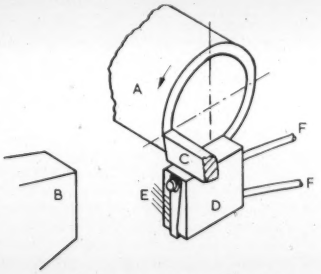


Fig. 2. Experimental arrangement.
A, work; B, camera; C, tool; D, furnace; E, thermocouple leads; FF′, cables from transformer.

in which the tube, 2 in. in outside diameter and with an $\frac{1}{8}$ in. thick wall, was machined in a lathe with a cutting tool slightly wider than the wall of the tube.

A plate camera having an $f/4{\cdot}5$ lens was focused on the side of the tool and the outside of the tube. A filter was fitted to the camera lens to exclude visible light and the camera encased in aluminium foil to prevent infra-red radiation from penetrating the bellows. It was also found necessary to erect screens round the apparatus to prevent the reflection of daylight from metallic surfaces in the vicinity of the apparatus.

A calibration furnace was arranged immediately below the tool so that the furnace and the tool, chip and workpiece could be photographed simultaneously. The furnace consisted basically of a tapered strip of soft copper–nickel alloy through which a heavy electric current was passed. This gave a linear variation in electric heating along the strip and an almost linear temperature distribution along its length. The temperature of the strip was measured at eight points along its length by chromel–alumel thermocouples folded inside the furnace strip and brazed (Fig. 3). With $\frac{1}{4}$ v applied to the terminals the temperature distribution along the strip ranged, approximately, from 300° c to 800° c. A shield on the furnace protected the strip from draughts caused by motion of the lathe chuck and prevented chips from damaging the thermocouple leads.

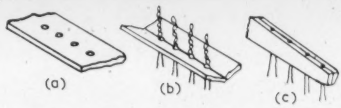


Fig. 3. Method of insertion of thermocouples in furnace strip.
(*a*) Holes drilled in strip.
(*b*) Ends of thermocouple leads twisted, brazed and inserted in holes.
(*c*) Strip folded along axis then cut to shape.

The fastest infra-red photographic plates available were Kodak I.R.E.R. fine-grain plates, which were sensitive down to a light wavelength of $0{\cdot}7\ \mu$. Measurements were made with these plates of the relationships between exposure time, the temperature of the surface photographed and the resulting optical density of the plates (Fig. 4). These results were

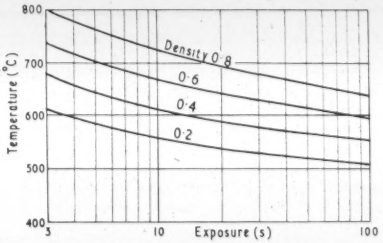


Fig. 4. Sensitivity of Kodak I.R.E.R. plates at $f/4{\cdot}5$.

obtained by photographing the calibration furnace using various exposure times and measuring the resulting distribution of optical density of the plate with a microphotometer. The curves indicate that for a given exposure the useful range of the plates was approximately 200 deg c, the range increasing at shorter exposure times; they also show that the lower limit of temperature measurement was approximately 500° c for an exposure of about 15 s.

With the workpiece initially at room temperature it was expected that, with the cutting conditions employed in these experiments, the shear zone temperature would be in the range 200° c to 300° c. This was below the useful temperature range of the photographic plates and it was therefore decided to pre-heat the workpiece and raise its initial temperature to about 350° c. It was found that this pre-heating of the workpiece had the additional advantage of improving the cutting conditions by the thermal softening of the material.

The radiating surfaces of the workpiece, tool and calibration furnace strip were given a thin coating of lamp black to ensure that they would have similar values of emissivity. As the technique was one of comparing, by photographic means, the temperatures of two objects, then, provided the emissivities of the two objects were similar, the actual value of emissivity was unimportant.

After pre-heating the workpiece and obtaining steady cutting conditions with no built-up edge fragments detectable at ×100 magnification either on the tool cutting edge or the underside of the chip, the photographic plate was given an exposure of approximately 15 s. The actual exposure time was not critical.

An example of an infra-red photograph obtained under the above conditions is given in Fig. 5. The distribution of optical density of the plate was measured by means of an automatic recording microdensitometer fitted with a circular aperture of 0·01 mm diameter.

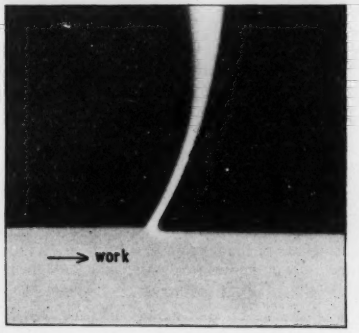


Fig. 5. Infra-red picture of orthogonal cutting process.

The distribution of optical density along the centre line of the picture of the calibration furnace was first obtained (Fig. 6) and the recording compared with a graph of the actual temperature distribution along the furnace strip; this enabled further records of optical density taken from the same plate to be interpreted in terms of temperature.

The photograph of the workpiece, chip and tool was then examined by obtaining records of optical density distribution along a series of lines spaced at 0·005 in. intervals. These readings, together with the calibration relationship obtained from the picture of the furnace, enabled the isotherm pattern to be drawn to a large scale (Fig. 7).

A typical record of optical density distribution along a path such as AB in Fig. 7 is shown in Fig. 8. The temperature distribution along AB shows the variation in temperature to which a particle of the work material was subjected as it passed through the shear zone into the chip. Twenty such records were used to construct the isotherm pattern of Fig. 7.

The points on the records defined by sudden changes in optical density indicated the boundaries of the workpiece, chip and tool, and when plotted enabled the geometric picture of the cutting process to be accurately superimposed on the isotherm patterns.

4. Discussion of results

The picture of the temperature distribution in the workpiece, chip and tool obtained by using the technique described in this paper was representative of the average temperature distribution occurring during a period of about 15 s. As both the cutting tool and the camera were rigidly secured to the lathe saddle, the temperature distribution for the cutting tool would be expected to be reasonably accurate. The results for the workpiece and chip, however, would have been affected by any change in the cutting conditions whilst the plate was exposed. Examination of the chip produced during the test showed that the chip thickness did not vary by more than 4%. This would not have affected the results significantly except those near the free surface of the chip.

The accuracy of the results depended also on the uniformity of development of the photographic plate. To give even development of the plate surface it was brushed during processing and, under these conditions, the variation in density may be 4% from the centre to the edge (Loftus 1957).

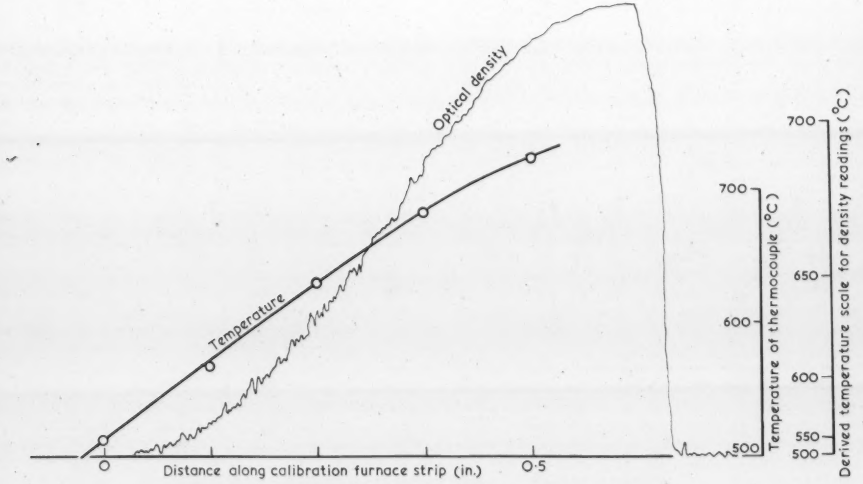


Fig. 6. Microdensitometer record along picture of calibration furnace.

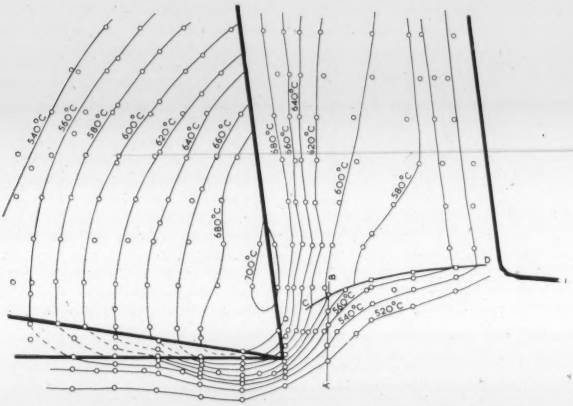


Fig. 7. Temperature distribution in shear zone, chip and tool during orthogonal cutting.

—○— = isotherms.
—□— = shear zone boundary.

Cutting conditions:
work material, 0·14 carbon steel (bright drawn)
cutting speed, 111 ft/min
undeformed chip thickness, 0·0119 in.
deformed chip thickness, 0·032 in.
tool material, Ardaloy carbide tip Grade S 200
tool rake angle, 10°
tool clearance angle, 7°.

As the pictures of the furnace and workpiece, chip and tool were small and were arranged equidistant from the edge of the plate, it was expected that the errors resulting from this effect would be negligible.

The temperatures obtained were those existing at the side surfaces of the workpiece, chip and tool and were estimated to be 2 or 3 deg C lower than the temperatures within the material, owing to the effect of radiation losses.

The effect of the differences in emissivity of the radiating surfaces was also estimated to be small. The density of the photographic plate gives a measure of the energy emitted by radiating surfaces, and this is proportional to the emissivity of the surface and to its temperature raised to the fourth power. Assuming that the emissivity of the calibration furnace strip surface was 1·0 and that of the work or tool surface 0·9, the apparent temperature would be approximately 0·97 of the real temperature, i.e. the result would be 3% low. In the experiments, all the surfaces were coated with lamp black and it is expected that the emissivity of the surfaces would not vary as much as the example quoted above.

This paper describes a technique with which it is hoped to examine, in detail, the temperatures existing in the workpiece, chip and tool in a variety of cutting conditions and then to compare the results with those obtained by analytical methods. A few tentative comments may be made, however, on the results presented here.

The position of maximum temperature at the tool rake face, some distance from the cutting edge, qualitatively agrees with the work of Rapier (1954) and Trent (1952) though the magnitude of this temperature (about 100 deg C above the average chip temperature) is not as high as that predicted using Rapier's analysis for these conditions (250 deg C above average chip temperature). His assumption of a plane heat source at the chip–tool interface would lead to an overestimation of maximum temperatures on the tool rake surface, because the heat source extends over a substantial width of chip.

In Fig. 8, the abrupt change in slope of the curve of optical density shows the point at which the material was leaving the

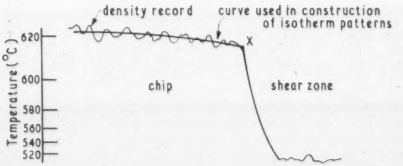


Fig. 8. Typical microdensitometer record through shear zone showing change in slope at X, a point on the shear zone boundary.

area of heat generation, and thus marks one boundary of the deformation or shear zone. Similar characteristics were obtained with all the records taken through the shear zone and the points at which the change in slope occurred are plotted in Fig. 7. The curve CD, through these points, therefore defines one boundary of the shear zone. The temperature distribution along this boundary is plotted in Fig. 9 and is qualitatively similar to Reichenbach's (1958) result. The steep temperature rise near the tool point is due to the presence of the frictional heat source at the chip–tool interface. In previous theoretical work the presence of this second heat source has been neglected when determining the temperature distribution along the shear zone.

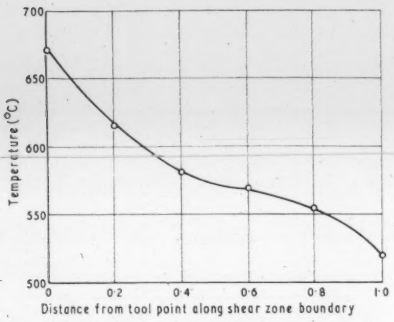


Fig. 9. Shear zone temperature distribution.

5. Conclusions

The photographic technique for the measurement of metal cutting temperatures has limitations. The sensitivity of the infra-red photographic plates is such that an exposure time of 10 or 15 s is required. The resulting photograph is an average of the temperature conditions prevailing during the exposure period and thus only stable cutting conditions may be examined by this technique; also high local temperatures, which may occur for short periods, will not be observed. In order to obtain complete temperature patterns for the workpiece, chip and tool it is necessary to pre-heat the workpiece between 350° C and 500° C, and also maintain cutting conditions which will not give maximum temperatures of more than 200 deg C above the initial workpiece temperature. This usually means that the cutting speed must be kept relatively low. It is also necessary to obtain conditions resulting in a smooth, continuous chip having a rectangular cross section, and the difficulty of obtaining such conditions increases as the cutting speed decreases.

The accuracy of absolute temperature measurement is reasonably good and, in view of the comments made in Section 4, it is expected that the values obtained may be 2 or 3% low. The accuracy of the isotherm patterns, however, should be better than this as the construction of the patterns depends on temperature differences, and the results show that differences of 20 deg C are well defined.

It is hoped that, by using the photographic technique, valuable information will be obtained regarding the shape and extent of the areas of deformation in orthogonal cutting.

The results presented indicate that future theoretical work on shear zone temperature distribution should include the effect of the frictional heat source at the chip–tool interface. Theoretical work on tool and chip temperatures should examine the effect of the deformation and resulting heat generation within the chip in the region of the tool–chip contact area. Work is being carried out by the author along these lines.

Measurement of the cutting forces simultaneously with temperature measurements will, in the future, enable direct comparisons of the measured temperatures with those obtained by analytical techniques.

Acknowledgments

The author wishes to thank the Governors of the Royal Technical College, Salford, for providing facilities for this work and Mr. A. W. J. Chisholm, Head of the Department of Mechanical Engineering, for his advice and encouragement. He also wishes to thank Mr. D. Glendinning for assistance with the photographic work.

References

CHAO, B. T., and TRIGGER, K. J., 1958, *Trans Amer. Soc. Mech. Engrs*, **80,** 311.

HERBERT, E. G., 1926, *Proc. Instn Mech. Engrs*, **1,** 289.

HOLLANDER, M. B., 1959, *American Society of Tool Engineers Research Report* No. 21 (*Detroit*).

HOLLANDER, M. B., and ENGLUND, J. E., 1957, *American Society of Tool Engineers Research Report* No. 7 (*Detroit*).

LOFTUS, A., 1957, *Spectrochemica Acta*, **9**, 216.

NAKAYAMA, K., 1956, *Bull. Fac. Engng Yokohama Nat. Univ.*, **5,** 1.

RAPIER, A. C., 1954, *Brit. J. Appl. Phys.*, **5,** 400.

REICHENBACH, G. S., 1958, *Trans Amer. Soc. Mech. Engrs*, **80,** 525.

TRENT, E. M., 1952, *Proc. Instn Mech. Engrs*, **166** (1), 64.

WEINER, J. H., 1955, *Trans Amer. Soc. Mech. Engrs*, **77,** 1331

Ultrasonic relaxation in polyatomic gases and vapours

by S. K. K. JATKAR, D.Sc., F.R.I.C., F.Inst.P., and D. D. DESHPANDE,* Department of Chemistry, University of Poona, India

MS. received 17*th October* 1960

Abstract

The paper describes a simple method of calculating ultrasonic relaxation frequencies in polyatomic gases. The mechanism of relaxation is explained as a resonance phenomenon by assuming a characteristic rotational oscillation frequency of gas molecules in ground state, due to the interaction field of neighbouring molecules. Using infra-red vibration and electronic frequencies given by the ionization potential, in the relation $f = (1/2\pi)\sqrt{(3h\nu\alpha^2/4r^6)}$ *the relaxation frequencies are evaluated for carbon dioxide, carbon disulphide, nitrous oxide, sulphur dioxide, ammonia, acetaldehyde, benzene, methane and chloromethanes, and are in close agreement with the experimental data. The method is absolute and does not involve calculation of parameters such as partition functions and collision probabilities.*

Introduction

Since the first observation by Pierce (1924, 1925) on the dispersion of ultrasonic velocity in gases with frequency, exhaustive data have been accumulated on several gases and vapours under varied experimental conditions such as pressure and temperature. The dispersion in gases is also accompanied by an anomalous absorption. Herzfeld and Rice (1928), Kneser (1931) and Rutgers (1933) developed the relaxation theory originally suggested by Jeans (1925). The theory states that the interchange of external energy of translation and internal energy of rotation or vibration of gas molecules no longer keeps pace with rapid pressure alterations produced by high-frequency sound waves. With the rapid adiabatic changes, the vibrational degrees of freedom no longer have time to attain equilibrium. With increase of frequency, therefore, the vibrational energy does not contribute towards the heat capacity. The lowering of heat capacity at higher sound frequencies gives rise to an increase in velocity. The theory assumed the loss of the whole vibrational contribution in a single step and the rotational energy term was included in the external energy. Bourgin (1929) suggested multiple transitions for both rotational and vibrational energy levels.

Schwartz, Slawsky and Herzfeld (1952) calculated the relaxation times quantum-mechanically, by evaluating the probability of vibrational de-excitation for typical encounters between pairs of molecules. The results for vibrational relaxation times thus calculated for several gases and vapours have been given by Schwartz and Herzfeld (1954), Tanzcos (1956) and Dickens and Linnett (1957). The order of magnitude of calculated relaxation times roughly compares with the experimental data. (See Table 1, column 5.)

* Now at the Indian Institute of Technology, Bombay.

In the present paper, the relaxation frequencies for polyatomic gases and vapours have been calculated by using the Lennard-Jones potential and London forces.

2. Intermolecular forces and ultrasonic dispersion

The ultrasonic dispersion and anomalous absorption occurs in the low-frequency region of 20 to 5000 kc/s. Richardson (1934) and Jatkar (1939) indicated that molecules showing high absorption and dispersion possess resonating structures in the majority of gases. Some non-resonating molecules like carbon tetrachloride, chloroform, methylene chloride, etc., also have been found to show absorption and dispersion. The most important property of these molecules is that they are optically anisotropic. Traces of impurities alter the dispersion region to a great extent (cf. Knudsen and Fricke (1940)). The change in pressure also shifts proportionally the dispersion region, a fact made use of by most of the investigators.

The ultrasonic absorption band shown by the gases is very broad. This may be due to a spectrum of relaxation times unresolvable owing to their close situation. In the present work, such an occurrence of a number of relaxation frequencies very close to each other has actually been observed.

Consider an anisotropic molecule like carbon dioxide. It will have a minimum two degrees of freedom of rotation, the rotations being around the axis through the centre of gravity of the molecule.

All the molecules are not, however, in the same energy state. According to the Boltzmann's distribution, the number of molecules in a particular energy level is given by $n_i = n_0 \exp(-U_i/kT)$, where n_0 is the total number of molecules and the other terms carry the usual significance. The number of molecules in the ground state will be $(n_0 - \Sigma n_i)$. These molecules in the ground state will have no degree of freedom for rotation. If two such molecules in the ground state approach each other, after collision they will undergo mechanical rotational oscillation. One can postulate, therefore, that there will be a natural frequency of these rotational oscillations. Ordinarily these rotational oscillations will have a small amplitude, but they will be excited by the propagation of sound waves if the wave period corresponds to the natural frequency of the rotational oscillations. Under such conditions a resonant absorption can take place, and as the amplitude of rotational oscillations will be maximum under the excited condition, the hindrance in rotation will give rise to an ineffective transfer of energy from translation to rotation and from rotation to vibration. This causes a partial knocking-out of vibrational heat capacity, giving rise to the dispersion of velocity at higher sound frequencies.

Since these rotational oscillations of molecules are governed

by the interaction field of neighbouring molecules, one can write, for the natural frequency of rotational oscillations,

$$f = (1/2\pi)\sqrt{(C/I)} \qquad (1)$$

where I is the moment of inertia of the molecule and the C is the couple acting on the molecules. At resonance, the ultrasonic relaxation frequencies should be identical with these rotational oscillation frequencies.

In order to evaluate these frequencies there has been substituted for C: (i) the interaction energy between a pair of molecules given by the Lennard-Jones potential function; (ii) interaction energy given by London dispersion forces; and (iii) the total attraction given by dispersion forces, orientation and induction effects.

The Lennard-Jones potential has been evaluated from $\phi(r) = 4\epsilon\{(\sigma/r)^{12} - (\sigma/r)^6\}$. The parameters used in the calculation given in Table 1 are taken from Hirschfelder, Curtiss and Bird (1954). The values of f calculated from $f = (1/2\pi)\sqrt{\{\phi(r)/I\}}$ are given in Table 1 and compared with the experimental values and the calculated values of Schwartz, Slawsky and Herzfeld (1952). Especially in carbon dioxide, the agreement between calculated and experimental results is better than in Herzfeld's method.

Table 1. *Lennard-Jones potential and ultrasonic dispersion*

Substance	ϵ/k (° K)	σ ($\times 10^8$ cm)	I ($\times 10^{40}$ g cm^2)	f Calculated (kc/s)	f Observed (kc/s)	
Carbon dioxide	190	3·996	70·6	83 362·3	17, 37 102	(*b*) (*c*)
Nitrous oxide	220	3·879	66·0	176	155	(*d*)
Carbon disulphide	448	4·438	253	425	380	(*d*)
Methane	143	3·832	5·3	259 55	125	(*f*) (*e*)
Carbon tetrachloride	471	5·528	487·6	2104 2600	4200	(*g*) (*e*)
Methyl chloride	414	4·044	5·5 63·6	2400 770 930	1130 840	(*h*) (*g*) (*e*)
Methylene chloride	398	4·783	26·2 252·6	4700 1500 120	2270	(*h*) (*e*)
Chloroform	437	5·134	259	3000 3400	9500	(*g*) (*e*)
Sulphur dioxide	252	4·290	12·3 73·2	1293 574	1040 2000	(*d*) (*i*)

(*a*) Schwartz, Slawsky and Herzfeld (1952) (extrapolated from their data on the probability of complex excitation).
(*b*) Pielemeier (1943).
(*c*) Buschmann and Schäfer (1941).
(*d*) Fricke (1940).
(*e*) Tanzcos (1956).
(*f*) Eucken and Aybar (1940).
(*g*) Fogg, Hanks and Lambert (1953).
(*h*) Sette, Busala and Hubbard (1955).
(*i*) Lambert and Salter (1957).

By substituting London's dispersion energy for C in Eqn (1) the rotational oscillation frequencies have been evaluated for a number of gases and vapours. These results have been given in Table 2. The Eqn (1) reduces to:

$$f = (1/2\pi)\sqrt{\{(3h\nu\alpha^2)/4r^6 I\}} \qquad (2)$$

where α is the polarizability of molecules, r is the intermolecular distance and ν is the electronic frequency, such that $h\nu$ is the ionization potential. In addition to this electronic frequency, a series of infra-red vibration frequencies have also been substituted for ν. The results given in Table 2 indicate that while the rotational oscillation frequencies calculated by using infra-red frequencies represent the onset of an S-shaped dispersion curve (cf. Figs 1–3), the

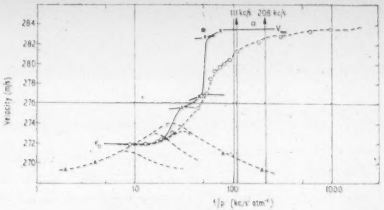


Fig. 1. Dispersion and absorption versus (f/p) in carbon dioxide at 30° c.
×—× = theoretical curve.
○—○ = velocity data, by Richards and Reid (1934).
△—△ = absorption, by Fricke (1940).
□ = Overbeck and Kendall (1941).
⊗ = Sujir (1948).

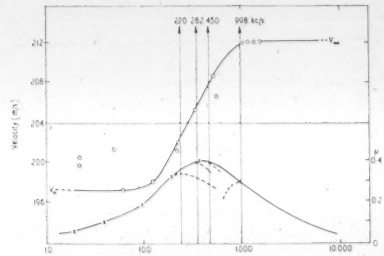


Fig. 2. Dispersion and absorption versus (f/p) in carbon disulphide at 30° c.
○—○ = velocity data, by Richards and Reid (1934).
×—× = absorption, by Fricke (1940).

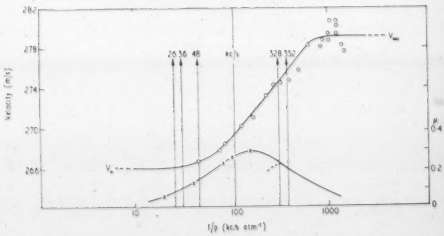


Fig. 3. Dispersion and absorption versus (f/p) in nitrous oxide at 20° c.
○—○ = velocity data by Richardson and Railston (1935).
×—× = absorption, by Fricke (1940).

value of f calculated from electronic frequency lies near the inflection point, and the upper region of the dispersion curve is limited by the value of f calculated from the Lennard-Jones potential.

The fact that the infra-red frequencies are also responsible for ultrasonic absorption at low frequencies signifies that

Table 2. *London forces and ultrasonic dispersion*

Substance	$r \times 10^{6}$ (*cm*)	$\alpha \times 10^{25}$ (cm^3)	$I \times 10^{40}$ ($g\,cm^2$)	ν (cm^{-1})	f_m (*kc/s*) *Calculated*	*Observed*
Carbon dioxide	5·68	26·5	70·6	667	14	
				1 285	19	17,37 (*a*)
				2 349	25	24 (*b*)
				111 092 (*e*)	172	102 (*c*)
Carbon disulphide	4·713	87·4	253	397	110	
				657	141	
				1 523	215	
				81 205 (*e*)	499	380 (*b*)
Nitrous oxide	6·1	30·00	66·0	582	13	
				1 285	18	
				2 223	24	
				104 000 (*e*)	164	153 (*b*)
Sulphur dioxide	3·53	37·2	12·3	519	107	
				1 151	159	
				1 361	173	
				99 411	1 477	1 040 (*b*)
				B	11 400	2 000 (*d*)
			85·5	519	59	
				1 151	89	
				1 361	98	
				99 411	829	1 040 (*b*)
				B	4 300	2 000 (*d*)
Ammonia	6·88	22·6	2·82	931	25	
				1 627	32	
				3 410	47	
				92 700 (*e*)	255	
				B	1 650	
			4·698	931	20	
				1 627	26	
				3 410	37	
				92 700 (*e*)	195	
				B	1 300	896 (*e*)
Acetaldehyde	2·5	46·72	22·36	512	278	
				1 000	388	600 } (*f*)
				3 000	672	2 800 }
Methane	7·532	26·0	5·298	1 306	19	
				1 526	20	
				2 914	27	
				3 020	28	
				104 647 (*e*)	520	125 (*g*)
Carbon tetrachloride	2·175	105	487·6	218	133	
				458	192	
				797	253	
				1 546	353	4 200 (*h*)
				92 402	2 700	
Methyl chloride	4·486	45·6	5·492	90 872 (*e*)	1 262	1 130 (*g*)
	4·486	45·6	63·6	90 872 (*e*)	371	840 (*h*)
Methylene chloride	3·25	64·8	26·21	91 436 (*e*)	2 165	2 270 (*i*)
					697	
Chloroform	2·763	82·3	259·9	91 999 (*e*)	1 426	9 500 (*h*)
				B		
Benzene	2·665	103	147·6	671	179	
				1 037	223	
				1 485	267	
				3 099	385	
				74 478 (*e*)	2 378	2 000 (*j*)

The values of ν denoted by (*e*) are electronic frequencies evaluated from ionization potentials taken from the recent data of Watnabe (1957). The rest of the values are infra-red frequencies given by Herzberg (1945); *B* denotes the sum of total interactions of dispersion, induction and orientation.

(*a*) Pielemeier (1943).
(*b*) Fricke (1940).
(*c*) Buschmann and Schäfer (1941).
(*d*) Lambert and Salter (1957).
(*e*) Özdogan (1950).
(*f*) Alexander and Lambert (1942).
(*g*) Eucken and Aybar (1940).
(*h*) Fogg, Hanks and Lambert (1953).
(*i*) Sette, Busala and Hubbard (1955).
(*j*) Cheng (1951).

they produce small intermolecular forces. In a separate work (unpublished) the authors have shown that the addition of infra-red frequencies to the electronic frequency brings the value of the Van der Waals constant b nearer the experimental value.

For polar molecules the values of rotational oscillation frequencies are also evaluated by using the total attractive force constant given by:

$$B = \{(3h\nu\alpha^2/4) + (2\mu^4/3kT) + 2\mu^2\alpha^2\}. \qquad (3)$$

An analogous treatment satisfactorily explains the dielectric absorption and relaxation in liquids, on the basis of rotational oscillation frequencies of the molecules. This work will be published elsewhere.

In the light of these rotational oscillation frequencies calculated as above, the mechanism of ultrasonic relaxation can be explained as follows. As previously suggested by Einstein (1920) in the case of a dissociating gas, it is assumed that the energy transfer takes place from translation to rotation and from rotation to vibration in steps. When the acoustic frequency is lower than the rotational oscillation frequency governed by the lowest infra-red vibration, the vibrational energy has sufficient time to absorb quanta of sound energy and the gas shows a normal heat capacity. As the acoustic frequency slowly approaches the value of the lowest rotational oscillation frequency, the energy transfer between the rotational level and the first vibrational level cannot keep pace with the rapid oscillations in pressure produced by the passage of sound waves. This means that the first vibrational level does not contribute towards the heat capacity, thus giving rise to a partial increase in velocity. When the sound frequency increases to the next value of the rotational oscillation frequency, the contribution of this mode of vibration to the heat capacity also disappears. A stepwise knocking-out of heat capacity is therefore expected, giving rise to a multiple transition due to each rotational oscillation frequency in all gases. An experimental resolution of such multiple transition is not always possible, unless the rotational oscillation frequencies are sufficiently separated from each other, and they will be so if there is a considerable difference between the infra-red vibration frequencies. Sette *et al.* (1955) and Dickens and Linnett (1957) have rightly observed that multiple transition can be detected in those gases of which the first and the second infra-red vibrations differ considerably from each other. When the acoustic frequency is identical with that rotational oscillation frequency governed by the electronic frequency, the loss of entire vibrational heat capacity takes place at this stage. Since the predominant inter-molecular forces are due to electronic frequency, the maximum resonant absorption also takes place at this frequency. The main broad absorption peak is also expected to be composed of a series of minor peaks due to other rotational oscillation frequencies.

The polarizabilities used in the present calculations are the arithmetic mean of the parallel ($\alpha_{\|}$) and perpendicular ($\alpha_{\perp}$) polarizabilities of the molecules. However, it is reasonable to use only parallel polarizability of the molecule under consideration. Since the rotational oscillation frequency is directly proportional to α, a correction may be applied to the values of f given in Table 2, by multiplying by the ratio $(\alpha_{\|}/\alpha_{\perp})$. According to the data given by Hirschfelder, Curtiss and Bird (1954), the ratios $(\alpha_{\|}/\alpha_{\perp})$ are 1·55 for carbon dioxide, 1·72 for carbon disulphide, 1·62 for nitrous oxide, 1·48 for sulphur dioxide and nearly unity for the rest of the molecules under present considerations. The corrected value of f is in very good agreement with the experimental data, especially for carbon dioxide and nitrous oxide. The values of f_m (Table 2, column 6) are corrected in this way for carbon dioxide, carbon disulphide, nitrous oxide and sulphur dioxide.

The mean free path r used in the present work is calculated from the relation $r = 1/\sqrt{2}\pi n\sigma^2$ which is deduced on the assumption that molecular velocities are distributed according to the Maxwell's distribution law. The actual molecular velocity fluctuates considerably from the root mean square value. Thus the intermolecular distance between a pair of molecules also fluctuates considerably about the average value of the mean free path. Such variations in the molecular distance sensibly affect the relaxation frequencies, and are responsible for the broadening of the absorption maxima.

The shift of the ultrasonic dispersion region with pressure can be explained as being due to the change in the interaction field between the molecules. When the pressure is increased, the molecular distance r is diminished and hence the rotational oscillation frequency increases, resulting in the shift of the dispersion region to a higher frequency range.

The mechanism of molecular rotation assumed is confined only to those molecules librating in the anisotropic molecular field, and such molecules are obviously those whose rotational energy does not enable them to rotate freely, i.e. the dispersion is due to molecules in ground state. The effect of increase in temperature will be to reduce the number of molecules showing ultrasonic absorption and dispersion, which may be shifted to higher frequencies or disappear at the lower frequencies.

It has been observed that the values of rotational oscillation frequencies given by infra-red vibrations lie on the onset of the dispersion curve. By evaluating the Planck–Einstein functions for the contribution of vibrational heat capacity C_i, using these infra-red vibrations, and subtracting it successively from the total heat capacity, it is possible to construct a theoretical velocity–frequency dispersion curve. The dispersed velocity in carbon dioxide at 30° C has been calculated in this way using the relation

$$V^2\phi = \frac{RT}{M}\,\frac{C_P - C_i}{C_v - C_i} \qquad (4)$$

where ϕ is the correction factor for the ideal gas state given by

$$\phi = 1 - \frac{P}{P_c}\,\frac{T_c}{T}\left\{1 - 6\left(\frac{T_c}{T}\right)^2\right\}. \qquad (5)$$

The loss of vibrational contribution has been calculated by assuming that the deformation vibration vanishes first, then the symmetrical and lastly, its degeneracy. This mechanism avoids a steep rise in velocity at low frequencies (near 15 kc/s) and the theoretical curve is nearer to the experimental points. Since the calculated relaxation frequencies represent the midpoint (inflection) of the sigmoid curve, the theoretical curves can be superposed exactly on the experimental velocity curves, if the dispersed velocity values (calculated) are plotted at twice the value of calculated relaxation frequencies. It is interesting to note that the absorption peak at 24 kc/s observed by Fricke (1940) exactly coincides with the value of relaxation frequency due to an entire loss of vibrational contribution.

The results of these calculations are given in Table 3 and shown graphically in Fig. 1, together with the experimental data on velocity and absorption. Figs 2 and 3 show the

Table 3. *Dispersed velocity in carbon dioxide at* 30° C

Infra-red vibrations lost	*f* (kc/s)	C_s (cal mol^{-1} deg C^{-1})	*V* (m/s)
—	—	—	270·8
ν_1	18	0·159	271·6
$\nu_1 + \nu_2$	30	0·163	271·7
$\nu_1 + \nu_2 + \nu_3$	31	1·059	276·5
$\nu_1 + 2\nu_2 + \nu_3$	36	1·955	283·0

dispersion and absorption in carbon disulphide and nitrous oxide respectively. The vertical lines on the graphs indicate the dispersion frequencies *f* given in Table 2.

Summary and Conclusions

The ultrasonic relaxation in polyatomic gases observed in the low-frequency region has been explained as a resonance phenomenon. The relaxation frequencies for several gases and vapours have been evaluated by postulating a characteristic frequency of gas molecules in the mutual field of intermolecular interaction. The calculated relaxation frequencies using London forces show an exceedingly close agreement with the observed values. From the knowledge of simple molecular constants like moment of inertia, ionization potential, mean free path and the polarizability, the dispersion region of any gas or vapour that is expected to show a vibrational relaxation can be predicted.

Acknowledgments

This paper contains a part of the Ph.D. thesis submitted to the University of Poona in 1956 by one of the authors (D.D.D.).

References

ALEXANDER, E. A., and LAMBERT, J. D., 1942, *Proc. Roy. Soc. A*, **179,** 499.

BOURGIN, D. G., 1929, *Phil. Mag.* (7), **7,** 821.

BUSCHMANN, K. F., and SCHÄFER, K., 1941, *Z. Phys. Chem. B*, **50,** 73.

CHENG, L. M., 1951, *J. Chem. Phys.*, **19,** 693.

DICKENS, P. G., and LINNETT, F. W., 1957, *Proc. Roy. Soc. A*, **243,** 84.

EINSTEIN, A., 1920, *Sitz. Ber. Berliner Akad.*, 380.

EUCKEN, A., and AYBAR, S., 1940, *Z. Phys. Chem. B*, **46,** 195.

FOGG, P. T. G., HANKS, P. A., and LAMBERT, J. D., 1953, *Proc. Roy. Soc. A*, **219,** 490.

FRICKE, E. F., 1940, *J. Acoust. Soc. Amer.*, **12,** 245.

HERZBERG, G., 1945, *Infra-red and Raman Spectra of Polyatomic Molecules* (New York: D. Van Nostrand Book Co. Inc.).

HERZFELD, K. F., and RICE, F. C., 1928, *Phys. Rev.* (2), **31,** 691.

HIRSCHFELDER, J. O., CURTISS, C. F., and BIRD, R. B., 1954, *Molecular Theory of Liquids and Gases* (New York: John Wiley and Sons Inc.).

JATKAR, S. K. K., 1939, *Indian J. Phys.*, **13,** 445.

JEANS, J. H., 1925, *The Thermodynamical Theory of Gases* (Cambridge University Press).

KNESER, H. O., 1931, *Ann. Phys.* (*Leipzig*) (5), **11,** 761, 777

KNUDSEN, U. O., and FRICKE, E. F., 1940, *J. Acoust. Soc. Amer.*, **12,** 255.

LAMBERT, J. D., and SALTER, R., 1957, *Proc. Roy. Soc. A*, **243,** 78.

LAXMINARAYANAN, D., 1950, Ph.D.Thesis (University of Madras).

OVERBECK, C. F., and KENDALL, H. C., 1941, *J. Acoust. Soc. Amer.*, **13,** 26.

OZDOGAN, I., 1950, *Istanbul, Fak. Mec. A*, **15,** 163.

PIELEMEIER, W. H., 1943, *J. Acoust. Soc. Amer.*, **13,** 32.

PIERCE, G. W., 1924, *Proc. Amer. Acad.* (*Boston*), **59,** 276.

—— 1925, *Proc. Amer. Acad.* (*Boston*), **60,** 271.

RICHARDS, W. J., and REID, F. A., 1934, *J. Chem. Phys.*, **2,** 193, 206.

RICHARDSON, E. G., 1934, *Proc. Roy. Soc. A*, **146,** 56.

RICHARDSON, E. G., and RAILSTON, 1935, *Proc. Phys. Soc.* (*London*), **47,** 533.

RUTGERS, A. F., 1933, *Ann. Phys.* (*Leipzig*) (5), **16,** 350.

SCHWARTZ, R. N., SLAWSKY, Z. L., and HERZFELD, K. F., 1952, *J. Chem. Phys.*, **20,** 1591.

SCHWARTZ, R. N., and HERZFELD, K. F., 1954, *J. Chem. Phys.*, **22,** 767.

SETTE, D., BUSALA, A., and HUBBARD, J. C., 1955, *J. Chem. Phys.*, **23,** 787.

SUJIR, R. J., 1948, M.Sc. Thesis (University of Bombay).

TANZCOS, F. I., 1956, *J. Chem. Phys.*, **25,** 439.

WATANABE, K., 1957, *J. Chem. Phys.*, **26,** 542.

Simplifying the Fourier integral by adequate specification

by C. G. MAYO, M.A., B.Sc., M.I.E.E., and J. W. HEAD, M.A., F.Inst.P., Research Department, B.B.C. Engineering Division, Kingswood Warren, Surrey

MS. received 2nd December 1960, in revised form 4th January 1961

Abstract

No quantity can be measured exactly, but only within a tolerance, however small; 'exact' measurement would require infinite information. Likewise conceptual ideas such as 'unit impulse' cannot be exactly realized, but they have realizable 'metrical' counterparts, which are approximations within a tolerance specified in advance.

Most of the difficulties associated with the Fourier integral vanish if the conceptual unit impulse (Dirac delta function with Dirac's limit closed), which plays an important part in the derivation of that integral, is replaced by a form in which the limiting technique is maintained. This form may be regarded as the output voltage resulting when a conceptual unit impulse of current flows in a unit 'pure' resistance shunted by residual capacitance. It can be represented as a function of time by the expression $\eta^{-1}e^{-t/\eta}H(t)$, where η, representing there siduale apacitance, is vanishingly small but not zero. It is the presence of this quantity η which removes all ambiguities and difficulties.

In the usual treatment of this question in terms of Jordan's lemma, a residual term is neglected which is estimated in terms of the size of a relevant integrand; if, however, the oscillatory nature of this integrand is also taken into account, the residual is found, as here shown, to be an absolute zero.

1. Introduction

No quantity can be measured exactly, but only within a specified tolerance, however small. Thus a resistor made to a tolerance of $\pm 20\%$ having a nominal value of 10 Ω might be expected to have an actual resistance of any value between about 8 Ω and 12 Ω. If the tolerance was reduced to $\pm 10\%$, $\pm 1\%$ or even $\pm 0{\cdot}1\%$, the possible variation of the actual resistance would be reduced and the cost of the resistor would be correspondingly increased, but it is not possible to make a resistor whose resistance is exactly 10 Ω. From the point of view of information theory, however, even if the resistance could be measured exactly, it could not be specified. This is made clear by writing the required value of the resistance as $10{\cdot}000\,000\ldots\,\Omega$; the specification that each figure after the decimal point is to be a zero requires $\log_2 10$ bits of information, and therefore $N \log_2 10$ bits are required for N zeros. This tends to infinity with N.

Likewise, there are many conceptual ideas useful in electromagnetic theory for which exact realization, or even unambiguous and verifiable specification, is not possible, but there are realizable and specifiable 'metrical' counterparts which can resemble the original conceptual idea within an arbitrarily close tolerance. One of these conceptual ideas, which is of vital importance in the study of the Fourier integral, is the unit impulse (Dirac delta function). Now, although the understanding of electrical phenomena is often greatly helped by regarding an arbitrary waveform $f(t)$ as made up of spectral components, there are considerable mathematical difficulties associated with the Fourier integral, which is explicitly or implicitly involved in any analysis in terms of spectral components. Such analysis can be carried out by first expressing the arbitrary waveform as a sum or integral of impulse functions, and then analysing each impulse into spectral components. The difficulties associated with the Fourier integral arise because the impulse function used is wrongly specified. If this conceptual impulse is replaced by a specifiable and realizable counterpart, all the difficulties vanish, as will be shown in § 2.

This realizable metrical counterpart of the unit impulse could be derived in a purely mathematical manner, but it is more illuminating not to disdain the help of intuition based upon physical considerations.* The excitation will be regarded as the voltage resulting when a conceptual unit impulse of current is delivered through a unit pure resistance. If, however, the unit pure resistance has associated with it a residual shunt capacitance η, the output voltage will be slightly but significantly different from a conceptual unit voltage impulse. This output voltage is the required 'metrical' unit voltage impulse, in contrast to the 'conceptual' unit current impulse which gave rise to it. The remarkable convergence of the modified Fourier integral based upon this metrical unit impulse is brought about not so much by the smallness of the integrand as by its rapid oscillation (for large values of the frequency variable of integration); this convergence is discussed in § 3.

Another form of inadequate specification based upon misuse of conceptual ideas (which is likely to cause dangerous confusion when, as here, fundamental ideas and principles are under consideration) is a strictly common-sense statement of the form 'The voltage between these terminals is 230 v'. Literally interpreted, this means that the voltage has been 230 v since the beginning of creation, and will remain so until the end of eternity, whereas the idea in the mind of anyone making such a statement is usually that the voltage is 230 v for all relevant times, and that for other times the value of the voltage does not matter. Now the proper value by means of which to express a quantity which does not matter is zero. The voltage in question could therefore have been better specified as $230\{H(t-a) - H(t-b)\}$ volts, where $H(x)$ means Heaviside's unit-step function, zero for negative x and unity for positive x. For this expression

* For example, van der Pol (1937) gives a very succinct discussion of shot noise, and shows how analysis in terms of the conceptual impulse leads to infinite noise power, which becomes finite when stray capacitance is taken into account.

means exactly what it says and is a complete specification for all values of the relevant variable t. The voltage in question was zero (or irrelevant) before time a, and zero (or irrelevant) after time b, but was 230 v at times between a and b, within which interval the experiment was begun and completed. More generally, a properly specified function, say $f(t)$ of time t, must be defined for all possible values of the argument (here t). Usually it will be like an island in the middle of a (one-dimensional) ocean: only different from zero for a small range of values of the argument. Thus in the example quoted the 'island' is only above the 'ocean' for times between a and b, both near the present time; in this case the island has steep 'cliffs' at times a and b, but this is not essential for adequate specification.

When the specification contains Heaviside unit-step functions, there will usually be an obvious reason for their presence, such as the start or finish of the experiment at particular times. In astronomical problems, however, the 'state equations' for the universe must necessarily be 'stationary'; otherwise special acts of creation or destruction have to be postulated. The limited epoch of observable phenomena (which are the subject matter of science) does not conflict with the stationary properties postulated for the universe, since the universal presence of entropy (which in the physical world resembles forgetting in the mental world) enforces this apparent limitation of epoch from the point of view of mankind.

2. 'Metrical' modification of the Fourier integral

For the present investigation, perhaps the most useful way to derive the Fourier integral is to start with the 'spotting' identity

$$F(t) = \int_{-\infty}^{\infty} \delta(t-\lambda)F(\lambda)d\lambda \tag{1}$$

in which $\delta(x)$ represents the conceptual or idealized Dirac function described by the equations

$$\delta(x) = 0 \ (x \neq 0) \quad \int_{-\infty}^{\infty} \delta(x)dx = 1. \tag{2}$$

Dirac observed that in general $\delta(x)$ could be used to obtain the same results as if a more cumbersome limiting form (as given, for example, by van der Pol 1937) were used instead. Eqn (1) is a truism if $\delta(t-\lambda)$ means the conceptual impulse specified by Eqn (2); to make it more than a truism, $\delta(t-\lambda)$ must be expressed in terms of other functions. In particular, if $\delta(x)$ is expressed in terms of sinusoids, and thus replaced in Eqn (1) by $\delta_F(x)$ where

$$\delta_F(x) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} e^{zx}dz = \frac{1}{\pi}\int_0^{\infty} \cos\,\omega x d\omega = \left(\frac{\sin\,\omega_0 x}{\pi x}\right)_{\omega_0 \to \infty} \tag{3}$$

we obtain the classical Fourier expression together with all the associated classical 'ifs and buts'. We should note first that the last member of Eqn (3) does not tend to any limit as ω_0 tends to infinity. This is only a symptom that all is not well; the essential point about Eqn (3) is that its formulation or specification is defective.

Van der Pol (1937) gives a number of expressions which approach $\delta(x)$ when some parameter tends to a limiting value. To these might be added the expression

$$\delta_m(x) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{e^{zx}dz}{1+\eta z} \tag{4}$$

which reduces to Eqn (3) if η is made zero, but not if η tends to zero. The expression (4) is suggested as worth consideration by the idea mentioned above that a voltage impulse might be regarded as the output when a conceptual unit current impulse is applied to a pure unit resistance with which is associated some parallel capacitance (represented by the normalized, dimensionless number η) however small. But it is shown below that $\delta_m(x)$ depends critically upon the sign of η however small η may be; thus

$$\delta_m(x) = \frac{1}{\eta}e^{-x/\eta}H(x) \quad \eta > 0 \tag{5}$$

$$\delta_m(x) = -\frac{1}{\eta}e^{-x/\eta}H(-x) \quad \eta < 0 \tag{6}$$

where $H(x)$ is Heaviside's unit-step function, zero for negative x and unity for positive x. Now Eqns (5) and (6) are radically different however small η may be, and this indicates that replacing η by zero is inadequate. For positive η, $\delta_m(x)$ is zero for negative x, large and positive for x small compared with η, and small and positive for x greater than say 5η. For negative η, however, $\delta_m(x)$ is the reflection of this about the line $x = 0$. For any η different from zero, the integral (4) can be exactly evaluated, that is to say, it is possible to find a value z_0 or $i\omega_0$ of the variable z of integration such that increasing z/i beyond ω_0 makes a contribution to the integral which is less than any assigned quantity. If η is zero, however, we saw from Eqn (3) that this cannot be done, and the integral remains a function of ω_0. The significant difference between Eqns (5) and (6) is that, for positive η, $\delta_m(x)$ is zero for negative x whereas, for negative η, $\delta_m(x)$ is zero for positive x. This is related to the existence of entropy and the ability to distinguish past from future. The remarkable convergence of the integral in Eqn (4) is worth further investigation, which is given in the next section.

If in Eqn (1) $\delta(t-\lambda)$ is replaced by $\delta_m(t-\lambda)$ derived from Eqn (4) (with η positive) the usual form of the Fourier integral is replaced by a 'metrical' form, namely

$$F(t) = \int_{-\infty}^{\infty} F(\lambda)\left(\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{e^{z(t-\lambda)}dz}{1+\eta z}\right)d\lambda. \tag{7}$$

If η is regarded as arbitrarily small, this can be reduced to

$$F(t) = \lim_{\eta \to 0} \int_{-\infty}^{\infty} \frac{1}{\eta}e^{-(t-\lambda)/\eta}F(\lambda)H(t-\lambda)d\lambda \tag{8}$$

but it is more fruitful to reverse the order of integration. A sufficient condition to permit this is

$$\phi = \int_{-\infty}^{\infty} |F(t)|dt \text{ is finite} \tag{9}$$

and this is satisfied by any $F(t)$ which genuinely represents a conceivable state of affairs. When we use Eqn (9) to reverse the order of the limits, we obtain the 'metrical Fourier integral' in the form

$$F(t) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\left(\int_{-\infty}^{\infty} F(\lambda)e^{-\lambda z}d\lambda\right)\frac{e^{zt}dz}{1+\eta z} \tag{10}$$

and this integral has the great advantage that there are no contradictions or ambiguities, and only straightforward mathematical procedures are required. The only restriction on $F(t)$ is specified by Eqn (9) and this is a restriction which should be welcomed, since it is automatically satisfied by the

only functions with which a physicist or engineer is concerned, namely those which are capable of representing states of affairs.

3. The convergence of the 'metrical impulse' integral

In the last section it was noted that, although Eqns (5) and (6) could be derived by the method of residues from Eqn (4), the remarkable convergence of the 'metrical impulse' integral specified by Eqn (4) called for further investigation. This convergence is brought about not so much by the smallness of the integrand for extreme z as by the rapid oscillation. Adequate advantage is not taken of this rapid oscillation in the usual development in terms of Jordan's lemma. In the following development a form of Dirichlet's integral is used instead to obtain the necessary extra precision.

In Eqn (4), put $zx = i\theta$; it reduces to

$$\delta_m(x) = \pm \frac{1}{2\pi i\eta}\int_{-\infty}^{\infty} \frac{e^{i\theta}d\theta}{\theta - (ix/\eta)}. \tag{11}$$

The doubtful sign in Eqn (11) is that of x; its presence is due to the fact that the limits of integration for θ contain the factor x. The integral in Eqn (11) is now split up into an infinite series of similar integrals; a typical member of this series has the limits of integration from $(2r-1)\pi$ to $(2r+1)\pi$, and in this the variable of integration is changed from θ to $(\theta + 2r\pi)$. Eqn (11) is thus equivalent to

$$\delta_m(x) = \pm \frac{1}{2\pi i\eta}\int_{-\pi}^{\pi} \sum_{r=-\infty}^{\infty} \frac{e^{i\theta}d\theta}{\theta - (ix/\eta) + 2r\pi}. \tag{12}$$

But it is well known* that

$$\sum_{r=-\infty}^{\infty} \frac{1}{\phi - r\pi} = \frac{1}{\phi} - \sum_{r=1}^{\infty} \frac{2\phi}{r^2\pi^2 - \phi^2} = \cot\phi \tag{13}$$

so that (using sgn x to denote $+1(x>0)$ and $-1(x<0)$)

$$\delta_m(x) = \frac{\operatorname{sgn} x}{2\pi i\eta}\int_{-\pi}^{\pi} \tfrac{1}{2}e^{i\theta}d\theta \cot\tfrac{1}{2}\{\theta - (ix/\eta)\} \tag{14}$$

$$= \frac{\operatorname{sgn} x}{2\pi i\eta}\int_{-\pi}^{\pi} \tfrac{1}{2}ie^{i\theta}d\theta \frac{e^{i\theta} + e^{-x/\eta}}{e^{i\theta} - e^{-x/\eta}} \tag{15}$$

which, by the substitution $\xi = \exp(i\theta)$, reduces to

$$\delta_m(x) = \frac{\operatorname{sgn} x}{4\pi i\eta}\int_C \left\{1 + \frac{2e^{-x/\eta}}{\xi - e^{-x/\eta}}\right\}d\xi \tag{16}$$

C being the unit circle. Now if x/η is positive the integrand in Eqn (16) has one pole within the unit circle C, whereas if x/η is negative, this pole is outside the unit circle and contributes nothing to the integral. The residue at the pole is $2\exp(-x/\eta)$. By systematic arrangement of the various sign possibilities, we thus arrive at Eqns (5) and (6), and there is no associated residual quantity whatsoever.

This is in sharp contrast to the usual treatment in terms of Jordan's lemma, in which the integral (4) is treated as a contour integral, completed by means of the left or right hand great semicircle as appropriate. Eqns (5) and (6) now follow only upon the assumption that the contribution of this great semicircle can be neglected; if this is taken to be of large radius R, we have, on substituting $Re^{i\theta}$ for z

$$|\text{great-circle contribution}| = \frac{1}{2\pi i}\left|\int_{\pi/2}^{3\pi/2} \frac{e^{Rx(\cos\theta + i\sin\theta)}iRe^{i\theta}d\theta}{1 + \eta Re^{i\theta}}\right| < \frac{1}{2\pi\eta}\int_{\pi/2}^{3\pi/2} e^{Rx\cos\theta}d\theta.$$

* See, for example, Jolley (1925, p. 126, Eqn (425)).

Substituting $\theta = \phi + \pi/2$, and taking account of symmetry, this reduces to

$$|\text{great-circle contribution}| < \frac{1}{\pi\eta}\int_0^{\pi/2} e^{-Rx\sin\phi}d\phi.$$

Over the range of integration, $\sin\phi > 2\phi/\pi$, so that finally

$$|\text{great-circle contribution}| < \frac{1}{\pi\eta}\int_0^{\pi} e^{-2Rx\phi/\pi}d\phi = \frac{1}{2Rx\eta}. \tag{17}$$

In formulating inequality (17) we have only taken into account the maximum possible size of the integrand, and have completely ignored its oscillatory nature. By this procedure, the definite statement of Eqn (5) must be replaced by

$$\left|\delta_m(x) - \frac{1}{\eta}e^{-x/\eta}H(x)\right| < \frac{1}{2Rx\eta} \quad \eta > 0 \tag{18}$$

so that if we are considering the 'spotting identity', Eqn (1), when $\delta(t-\lambda)$ is replaced by $\delta_m(t-\lambda)$ and $F(t)$ is say, an exponential, we have to account for a residual term

$$X = \int_{-\infty}^{\infty} \frac{e^{-\alpha\lambda}}{2R(t-\lambda)\eta}d\lambda \tag{19}$$

instead of being able to proceed immediately to Eqn (8). In ordinary language, the radius R of the great semicircle is supposed to tend to infinity in such a way as to make X, which has an arbitrarily large exponential factor in the numerator of its integrand, tend to zero. There does not appear to be any satisfactory way of overcoming this difficulty without making use of the oscillatory nature of the integrand as was done in deriving Eqn (16), and it cannot be too strongly emphasized that in fact Eqns (5) and (6) are true without any residual whatsoever, so that the 'spotting integral', Eqn (1), can be applied universally on the assumption that $\delta(t-\lambda)$ is replaced by $\delta_m(t-\lambda)$ with η arbitrarily small, but not zero. The restriction that $F(\lambda)$ must be capable of representing a state of affairs so that ϕ in Eqn (9) is finite has already been noted; it was imposed so that the order of integration could be reversed to derive Eqn (10). For Eqn (8), however, it is only necessary that $F(\lambda)$ should not grow more rapidly than $\exp(\alpha\lambda)$ where α is finite.

4. Conclusions

If a voltage impulse is regarded as a conceptual current impulse flowing in a pure resistance shunted by residual stray capacitance, a satisfactory and realizable representation is obtained. This realization is in fact exactly represented by $\delta_m(t)$ in Eqn (5), although the usual analysis in terms of Jordan's lemma fails to make clear the extreme accuracy of the representation, because it does not take into account the oscillatory nature of the relevant integrand as well as its smallness.

Acknowledgment

The authors wish to express their thanks to the Director of Engineering of the British Broadcasting Corporation for permission to publish this article.

References

JOLLEY, L. B. W., 1925, *Summation of Series* (London: Chapman and Hall).

VAN DER POL, B., 1937, *J. Instn Elect. Engrs*, **81**, 381.

Study of electric breakdown of liquid dielectrics using Schlieren optical techniques

by B. FARAZMAND, M.Sc., Grad.I.E.E.,* Electrical Engineering Department, University of Birmingham

Communicated by G. T. Wright, Ph.D., A.Inst.P.; MS. received 15th July 1960, in revised form 19th December 1960

Abstract

The work is a contribution to the understanding of the mechanism of electric breakdown in liquid dielectrics. The liquid dielectric chosen throughout the experiments was n-hexane, because an extensive literature exists on the breakdown of this material. The applied voltage was a rectangular pulse of different durations supplied from a 125 kv, five-stage Marx–Goodlet impulse generator. A novel rotary multiple electrode containing eight pairs of electrodes was incorporated in a simple and efficient fractionation–filtration unit. A protection circuit, incorporating an inverter and a trigatron, limited the duration of breakdown current.

A series of photographs of the state of n-hexane under the applied electric field was taken by Schlieren technique, using a spark light-source and a still camera. A statistical survey of the results shows the formation of a region of very low refractive index at the cathode. It is suggested that this is due to streamers with a very large number of branches, which propagate towards the earthed electrode. The rate of growth increases with voltage and above a certain voltage, if enough time is allowed, the stepped streamers bridge the electrode gap and breakdown of the liquid results.

1. Introduction

PIONEERING work on the electrical breakdown of liquid dielectrics was mainly concerned with the conditions prevailing in electrical apparatus. These were, basically, breakdown due to secondary effects like the presence of fibres, chemical impurities, water particles and electrode effects (Hirano 1945, Lewis 1953, Watson and Higham 1953).

Present investigations have been largely associated with breakdown due to the physical properties of liquid dielectrics, and to achieve this aim there has been tremendous progress in experimental techniques. The secondary effects in most cases have been eliminated.

Attempts to correlate experimental results with theories of breakdown processes are made difficult by the fact that there has been a considerable dispersion in the results obtained by any one worker. The deficiency in the information available is in the actual breakdown mechanism. The main difficulty in this work is the very short time involved, which is of the order of microseconds.

The author has concentrated his investigations on the transient state of liquids subjected to an electric field, with the aim of obtaining direct knowledge of the mechanism. So that the physical nature of the dielectric liquid should be clearly known, *n*-hexane has been selected: this is an aliphatic hydrocarbon of simple molecular structure on which an extensive literature on breakdown exists. (Crowe 1954, Goodwin and Macfadyen 1953, Green 1955, 1956, Macfadyen 1955, Sharbaugh, Bragg and Crowe 1955, Sharbaugh, Cox, Crowe and Auer 1955.)

Pre-breakdown and post-breakdown states of the liquid were photographed by a Toepler–Schlieren system. This is a technique rendering visible, on a viewing screen or photographic film, small local variations in the refractive index of a transparent medium.

There were two light sources available, one a mercury vapour used for aligning the optical components and the other a Libbesart spark gap giving a flash of very short duration for photography (Prescott 1952). A high-speed cathode-ray oscilloscope in conjunction with a camera was used to record the voltage across the test gap.

2. Experimental apparatus and techniques

2.1. *The control circuit.* The electronic control system enabled the firing of a spark light source to be delayed for a definite time with respect to the front of a high-voltage rectangular pulse. The control circuit also governed the duration of the high-voltage pulses and, through the protection circuit, minimized the damage to the electrodes and the liquid dielectric after each breakdown.

Figure 1 shows a block diagram of the electrical circuit. A positive pulse from a remote push-button pulser was applied to a single-shot multivibrator producing a rectangular pulse. This pulse was applied to a cathode follower through a differentiating network and a trigger amplifier. The pulse from the cathode follower triggered the first stage of a soft valve delay unit, which had a two-fold purpose. It started the time base of a cathode-ray oscilloscope and, after a fixed delay, it triggered its second stage, producing three pulse signals.

The first signal triggered a 3 kv hydrogen thyratron whose output fired a five-stage Marx–Goodlet impulse generator. The second signal through a variable soft valve delay unit triggered a 7 kv hydrogen thyratron, which in turn fired the chopping gap of the impulse generator. This determined the duration of the impulse wave applied to the liquid dielectric. The third signal was applied to a two-stage soft valve delay unit whose output triggered a 3 kv hydrogen thyratron, which fired the spark light source. When the dielectric broke down a signal was generated which, through the 7 kv thyratron, fired the chopping gap and thus terminated the applied voltage.

2.2. *The impulse generator.* The high-voltage supply was generated by a 125 kv five-stage Marx–Goodlet impulse generator. This was built from 1 μF capacitors, capable of withstanding a charging voltage of 25 kv, in a single column

* Now at the National Iranian Oil Co., Tehran, Iran.

mounting. The circuit was designed to give an output voltage with a very short rise time and a long decay time. Consequently, a rectangular pulse could be obtained by chopping the voltage at a predetermined short time after its initiation. Both initiation and chopping were performed by trigatron units (Husbands and Higham 1951).

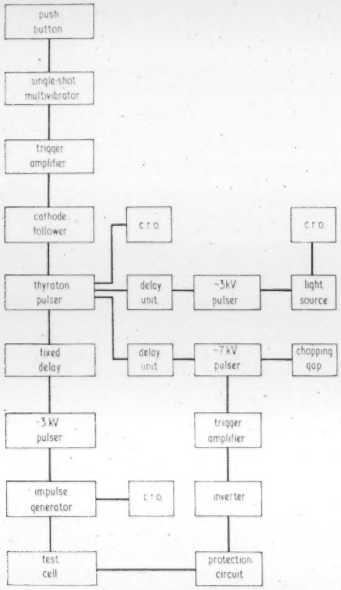


Fig. 1. Block diagram of electrical circuit.

When the liquid dielectric broke down a pulse was generated across a resistance in series with the test gap. This pulse, irrespective of its polarity, fired the chopping trigatron thus short-circuiting the test gap. The duration of post breakdown current through the test gap was reduced to less than 1 μs by the operation of the protection system.

2.3. *The rotary multiple electrode system.* In order to avoid the influence of electrode surface on the breakdown mechanism and to obtain a reliable and reproducible electric strength, it was necessary to use a pair of unblemished electrodes and a liquid sample free from impurities for each breakdown. With the Schlieren arrangement there was also the necessity of keeping the electrodes aligned in the optical axis.

To accomplish the above requirements a rotary multiple electrode system was designed, containing eight pairs of electrodes. The construction was simple and any electrode configuration—point, plane or sphere—could be used. The gap length was adjustable to any desired value and the electrodes could be mass produced.

2.4. *The fractionation–filtration system.* A circulatory fractionation–filtration system incorporating the test cell is shown in Fig. 2. The aim was to reduce the system to its simplest form, while retaining the main purpose of providing a chemically pure liquid dielectric in the test cell, free from fibres and particles.

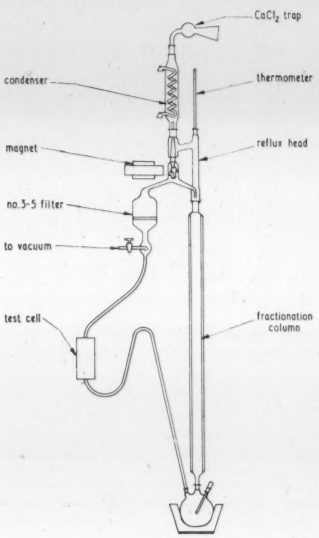


Fig. 2. The circulatory fractionation system.

The inlet and outlet to and from the test cell were flexible plastic tubes. This enabled the test cell to be aligned in the optical axis without disturbing the distillation system. The liquid was collected in the test cell by raising the outlet tube above the desired level of the liquid. By lowering the outlet tube the test cell could be drained. This had a great advantage over the previous systems used in that neither a tap nor vacuum drainage was needed. The liquid dielectric passed through a few cycles of fractionation and filtration before it was subjected to an electric field.

2.5. *The Toepler–Schlieren optical system.* Fig. 3 shows a schematic diagram of the optical system. A plane mirror M

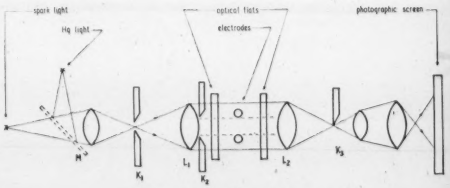


Fig. 3. The Schlieren system.

M, movable plane mirror; K_1, source slit; L_1, collimator head; K_2, limiting slit; L_2, Schlieren head; K_3, Schlieren diaphragm.

placed at 45° to the optical axis made it possible to select either a spark source or a continuous mercury light source, which later could be focused by a condenser lens on the source slit K_1. This enabled the optical components and test cell to be aligned very accurately.

The Schlieren diaphragm K_3 was placed at the focal plane of the Schlieren head. The sharp edge of K_3 was vertically adjustable to within 70 μ in. by a micrometer drive. The diaphragm could also rotate through 200° about the optical axis.

The source slit was placed at the focal point of the collimator head L_1. A slit K_2, close to L_1, limited the field of view, which was decided by the electrode configuration. The Schlieren head L_2 was a lens identical with L_1. None of the lenses in the optical system showed any internal Schlieren effects.

3. Results

The photographs presented in Figs 4 and 5 were selected from many experimental results, taken by the Schlieren technique to show the state of the liquid *n*-hexane under the influence of an applied electric field at increasing time intervals, from the instant of field application.

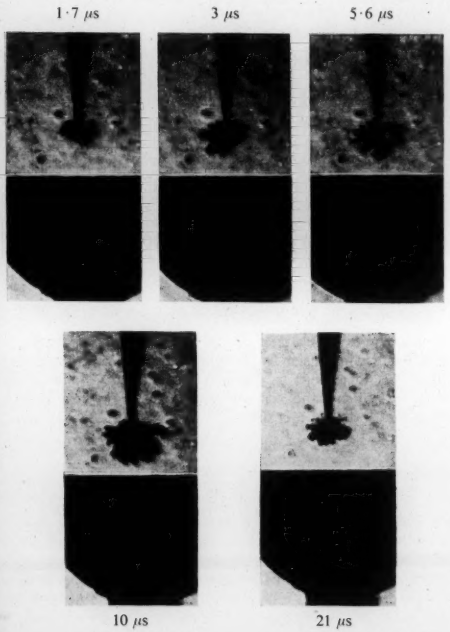


Fig. 4. Schlieren pictures of *n*-hexane at a field strength below breakdown. Time measured in microseconds from the instant of field application. Pulse duration, 10 μs; field strength, 300 kv/cm; gap length 1·5 mm; point–plane electrodes with print cathode.

The experimental results were obtained with the liquid at a temperature of 20–25° C and under normal pressure, without conditioning either the liquid or the electrodes.

The point–plane electrode configuration permitted large gap spacing without excessively high voltages. A gap length of 1–2 mm was used to give a clear picture of the inter-electrode space.

Rectangular pulse voltages were applied to the electrodes, starting from a voltage about half of the expected breakdown voltage and increasing in steps until the breakdown occurred. A recording of the high-voltage waveform was taken with each breakdown, thereby enabling the time of breakdown, to be assessed. In the electric field prior to breakdown a region of lower refractive index than that of the surrounding liquid was observed at the cathode. The extent of this region increased with time and with increase in field intensity.

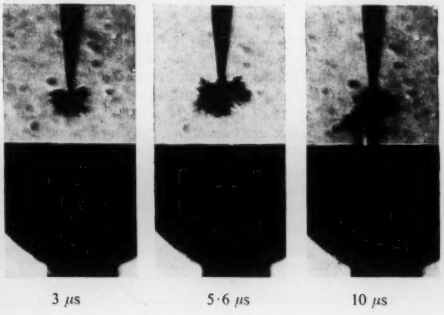


Fig. 5. Pre-breakdown Schlieren pictures of *n*-hexane leading to breakdown. Time measured in microseconds from the instant of field application. Pulse duration, 17·5 μs; breakdown after 12 μs; field strength, 310 kv cm; gap length, 1·5 mm; point–plane electrodes with print cathode.

It was also observed at field strengths as low as half the breakdown field strength. There was no visible change in the refractive index of the liquid when a positive instead of a negative rectangular pulse was applied to the point of the point–plane electrode system.

4. Discussion

The experimental evidence provides a clear qualitative picture of the behaviour of *n*-hexane under field strength at pre- and post-breakdown values. As the voltage increases from zero, positive ions accumulate at the converging field near the point cathode and set up an intense local field, resulting in cathode emission.

The electrons emitted from the cathode gain energy in the field and cause a streamer to start at the point, which may span as much as half the gap within 1 μs. The initiating streamer, which is followed by a series of stepped streamers, develops in the form of numerous branches which propagate until the earthed electrode is reached.

There is a remarkable similarity between the experimental results and the mechanism of natural lightning. Both phenomena appear to start with an electron avalanche rushing towards the plane from the negatively charged cloud or point; but the relative distance reached by the avalanche in the case of *n*-hexane seems to be much smaller than in the case of lightning.

A stepped streamer may penetrate to over 90% of the gap length without causing breakdown, but once it reaches the earthed electrode a heavy return stroke rushes to the point and breakdown of the liquid dielectric results.

The streamers grow with increase in voltage and time.

They have been recorded at a field strength as low as half the breakdown field.

Below breakdown field strength the growth of the streamer reaches a limit in size which does not change with increase in voltage duration. This is probably because the electric field intensity drops rapidly on leaving the point electrode to mid-gap values insufficient to cause ionization. An alternative reason may well be that loss of energy in exciting characteristic modes of vibration associated with carbon-hydrogen bonds is large enough to arrest any further propagation of the streamers (Sharbaugh, Cox, Crowe and Auer 1955, Lewis 1956, 1957).

Above a certain field strength, if enough time is allowed, the stepped streamers bridge the gap and breakdown occurs. This explains the increase in breakdown voltage with reduction of its duration and also explains the contradiction which exists between groups of investigators particularly as far as point–plane electrodes are concerned. With point electrodes at positive polarity no trace of any streamer has been recorded, but there is a need for further study before propagation by streamer can be ruled out.

The work presented is only an opening into a most fascinating field, and it is hoped that future studies will eventually provide a complete explanation of electric breakdown in liquid dielectrics under all conditions.

Acknowledgments

The author wishes to express his gratitude to Prof. D. G. Tucker for providing the facilities that made the findings in this paper possible.

He is also indebted to Dr. G. T. Wright for his advice and encouragement.

References

CROWE, R. W., 1954, *Annu. Rep. Conf. Elect. Insulation.*

GOODWIN, D, W., and MACFADYEN, K. A., 1953, *Proc. Phys. Soc.* (*London*) *B*, **66,** 85.

GREEN, W. B., 1955, *J. Appl. Phys.*, **26,** 1257.

—— 1956, *J. Appl. Phys.*, **27,** 921.

HIRANO, S., 1945, *J. Inst. Elect. Engrs* (*Japan*), **63,** 359.

HUSBANDS, A. S., and HIGHAM, J. B., 1951, *J. Sci. Instrum.*, **28,** 242.

LEWIS, T. J., 1953, *Proc. Instn Elect. Engrs*, Pt. IIA, **100,** 141.

—— 1956, *J. Appl. Phys.*, **27,** 645.

—— 1957, *J. Appl. Phys.*, **28,** 503.

MACFADYEN, K. A., 1955, *Brit. J. Appl. Phys.*, **6,** 1.

PRESCOTT, R., 1952, *Photogr. Engng*, **3,** 135.

SHARBAUGH, A. H., BRAGG, J. K., and CROWE, R. W., 1955, *J. Appl. Phys.*, **26,** 434.

SHARBAUGH, A. H., COX, E. B., CROWE, R. W., and AUER, P. L., 1955, *Annu. Rep. Conf. Elect. Insulation.*

WATSON, P. K., and HIGHAM, J. B., 1953, *Proc. Instn Elect. Engrs*, Pt. IIA, **100,** 168.

New books

Elektronische Hilfsmittel des Physikers. By W. GRUHLE. (Berlin, Göttingen, Heidelberg: Springer, 1960.) Pp. viii + 200. Price DM 29.60.

Dr. Gruhle's book, entitled *Electronic Aids for the Physicist*, is aimed at the experimental physicist who wants to design electronic measuring instruments in the course of his work in other branches of physics. It is 'neither a large textbook nor a collection of circuit diagrams', but provides the basis for learning to design one's own circuits by explaining the principle of the conventional solutions to specific problems in electronics.

The book begins with a discussion of circuit elements and goes on to deal with the generation of signals, a concept which includes a short section on radiation detectors. A chapter on the modification of signals discusses, *inter alia*, linear and non-linear amplification, stretching and delaying of pulses, and pulse shaping. There is a chapter on the combination of signals, and one on signal recording. The book concludes with a chapter on current and voltage stabilization.

This is an intelligent presentation of the fundamentals of electronics for the working physicist who is not a specialist in electronics. The task of the English-speaking reader is made somewhat easier by an English translation of the more ingenious inventions in German terminology. It is a pity that, by British standards, the price (around £3) seems rather high for a book of this size.

M. M. BLUHM

Modern university physics. By J. A. RICHARDS, F. W. SEARS, M. R. WEHR and M. W. ZEMANSKY. (London: Addison–Wesley, 1960.) Pp. xvi + 1006. Price 53*s*.

This book of about a thousand pages covers somewhat more than half the topics of an ordinary degree course in physics. A great deal of effort has clearly gone into finding simple derivations of important results, and the authors have not hesitated to confine themselves to a short qualitative account where a proper treatment might have been too difficult for a student taking an ordinary degree. The resulting level of presentation is probably below that required in most British Universities. Some suggestions for further reading would have been particularly valuable in a book of this type. There is a large number of simple examples.

A. HERZENBERG

Use of helium flush in the vacuum deposition of thin films

by G. N. SRIVASTAVA, Ph.D., and G. D. SCOTT, Ph.D., Department of Physics, University of Toronto, Canada

MS. received 21*st November* 1960

Abstract

Silver films deposited in a vacuum chamber previously flushed with an inert gas, helium, exhibit less aggregation as shown in electron micrographs and reduced ageing as determined from electrical conductivity, as compared with films produced under the usual high vacuum procedures. It is concluded that in general flushing the vacuum chamber with helium is of value in producing stable, compact and pure evaporated films.

One of the major reasons for the difference in structure between an evaporated film and the corresponding bulk material is the gas which is occluded in the film during its deposition. Two methods have been used to reduce the amount of occluded gas: high rates of deposition (Sennett and Scott 1950), and ultra-high vacuum (Bachmann *et al.* 1960). A third method, which is described here, is the use of an inert gas, helium, as the residual gas in the vacuum chamber.

A conventional 12 in. diameter evaporation unit with a 275 l./min oil diffusion pump was employed; it was equipped with a liquid air trap and had provision for discharge cleaning of the film substrate. The procedure for flushing the vacuum chamber to leave helium as the residual gas was as follows: the chamber was evacuated to a pressure of 10^{-4} mm Hg and then helium was admitted until the pressure rose to 10^{-1} mm Hg; the pressure was then allowed to fall to 10^{-4} mm Hg and helium again admitted. The flushing was repeated at least three times before the pressure was reduced to the final operating value of 10^{-6} mm Hg. It can be assumed that after several flushings with helium, the number of air molecules remaining in the chamber had been so reduced that the residual gas was essentially helium rather than air. The term *air* is loosely used here for the residual gas, which is usually composed of desorbed water and organic vapours.

To study the influence of the residual gas on film structure and properties, two sets of silver films were evaporated—one set under ordinary conditions, i.e. with air as the residual gas, and a second set when the vacuum chamber had been flushed with helium. Other factors influencing the structure of the film were kept the same. Silver films were prepared with different initial resistances in range from 10^2 to 10^7 ohms per square. One set of films were deposited in air, the second in helium. The films were formed on glass strips with electrical contacts as described by Chandra and Scott (1958). The initial resistance was noted immediately after deposition of the film and then readings were taken at regular intervals up to about 40 minutes after which time the resistance remained constant. The changing resistance during the period immediately following deposition is a measure of the ageing of the film. The thickness of the film was determined by the multiple interference technique.

The results showed that for all thicknesses the silver films formed in helium had a significantly lower final resistivity than those formed in air. The effect on the ageing of the silver films was quite marked. Fig. 1 shows the ageing of silver films with an initial resistance of 10^7 ohms. For films produced in ordinary vacuum the resistance increases about twenty times in 40 minutes, whereas for films deposited in helium the increase was only about ten times. Also shown is the effect on ageing made by the use of the liquid air trap. In the case of much thicker films a decrease in resistance occurs with ageing. For example, with silver films with an initial resistance of 10^2 ohms the final resistance is about 50% of the initial resistance when the film is deposited in air,

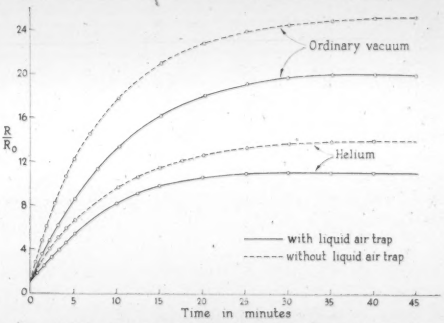


Fig. 1. The change in resistance of silver films with time for films evaporated under different conditions.

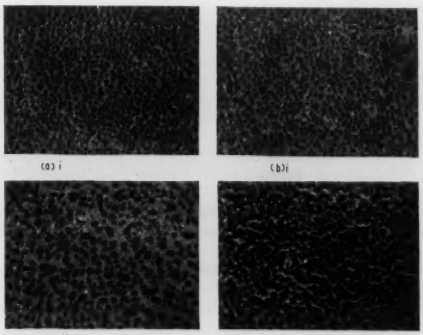


Fig. 2. Electron micrographs of silver films.
(*a*) evaporated under ordinary conditions: (i) 40 Å; (ii) 105 Å.
(*b*) evaporated with a helium flush: (i) 40 Å, (ii) 108 Å.

whereas the change is only to 60% when the film is deposited in helium. It can be stated that in general there is less ageing of the films, that is the final value of resistance is nearer the initial value, when the film is deposited with helium as the residual gas rather than air.

Electron micrographs were taken of silver films of different thicknesses deposited in air and in helium. Typical micrographs are reproduced in Fig. 2. For thinner films, i.e. about 40 Å 'mass' thickness, the difference in gross structure produced by the two conditions of evaporation is not significant. However, for thicker films, i.e. about 100 Å, the effect is quite marked. The film deposited in helium is much more continuous, showing only fissures, while the film deposited in air is highly aggregated.

In a series of experiments with evaporated iron films the 'helium flush' technique proved to be particularly useful. Whereas the electron diffraction pattern of iron films produced in ordinary vacuum always showed traces of oxides, the patterns of iron films deposited with helium as the residual gas were quite free of these impurity rings.

It is concluded that, together with other well recognized factors such as cleanliness of the substrate, high rate of deposition and low vacuum pressure, flushing with an inert gas can be of great value in producing stable, compact and pure evaporated films.

References

BACHMANN, L., ORR, W. H., RHODEN, T. N., and SIEGEL, B. M., 1960, *J. Appl. Phys.*, **31**, 1458.

CHANDRA, S., and SCOTT, G. D., 1958, *J. Sci. Instrum.*, **35**, 349.

SENNETT, R. S., and SCOTT, G. D., 1950, *J. Opt. Soc. Amer.*, **40**, 203.

Notes and comments

5th International Conference on Ionization Phenomena in Gases

The 5th International Conference, organized by Der Verband Deutscher Physikalischer Gesellschaften, Fachausschuss Plasma und Gasentladungsphysik, will be held in Munich, Germany, from 28th August–1st September 1961.

The following theoretical and experimental problems will be treated at the Conference: Fundamental processes in gas discharges: collision processes, radiation, transport phenomena, surface effects; Discharge types: glow discharge, arcs, pulse discharges, high frequency discharges; Diagnostic methods: spectroscopy, shock waves, microwaves; Applications. The Conference will be essentially a discussion meeting; preprints will be prepared from submitted manuscripts and distributed to participants prior to the Conference; the authors will then merely present a short summary of their work at the Conference and a Discussion Chairman will then attempt to initiate a lively discussion following presentation. Review papers and original papers, appended by additional remarks from the authors and contributions from the discussions, will be published after the Conference.

Further details are obtainable from V. Internationale Konferenz über Ionisationsphänomene in Gasen Sekretariat, München 1, Oskar-von-Miller-Ring 18, Germany.

British Electrical Power Convention, 1961

The Thirteenth British Electrical Power Convention will be held at Eastbourne from 12th–15th June 1961. The theme of the Convention will be "Electricity in the Prosperity and Welfare of the Nation". Papers will include: The British Electricity Transmission System; Some aspects of efficiency and economy in distribution; The electrical power industry in Canada; British electrical manufacture in the national economy. For further information apply to the Electrical Development Association, 2 Savoy Hill, London, W.C.2.

British Standards Institution

We have recently received notice of the following new publications of the British Standards Institution:

B.S. 3332: 1961—*White metal bearing alloy ingots*. Price 5*s*. B.S. 718: 1960—*Density hydrometers and specific gravity hydrometers*. Price 10*s*. B.S. 2634: Part 1: 1960—*Specification for roughness comparison specimens, ground flat and cylindrical types*. Price 4*s*. 6*d*. B.S. 1523: Section 2: 1960—*Glossary of terms used in automatic controlling and regulating systems, Process control*. Price 6*s*.

Selected Papers on Stress Analysis, presented at The Institute of Physics Stress Analysis Group Conference, Delft, 1959, Pp. 114 (London: Chapman and Hall Ltd. for the Institute of Physics, 1961).

This publication is now available to members at a reduced rate of 15*s*. per copy (cash with order) from The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1.

U.S. National Bureau of Standards

New Publications. *Research Highlights of the National Bureau of Standards, Annual Report* 1960, price 65 cents. *Precision Measurement and Calibration*, N.B.S. Handbook No. 77, issued in three volumes: Volume I, *Electricity and Electronics*, price 6 dollars; Volume II, *Heat and Mechanics*, price 6.75 dollars; Volume III, *Optics, Metrology and Radiation*, price 7 dollars. *Report of the International Commission on Radiological Units and Measurements* (*ICRU*) 1959, N.B.S. Handbook No. 78, superseding Handbook No. 62, price 65 cents. *Climatic Charts and Data of the Radio Refractive Index for the United States and the World*, N.B.S. Monograph No. 22, price 2 dollars. All these publications available from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C., U.S.A.

The colour-temperature of an incandescent helical filament

by J. E. GIBBS, Ph.D., A.Inst.P., and G. W. GORDON-SMITH, A.M.I.E.E., National Physical Laboratory, Teddington, Middlesex

MS. received 21st October 1960

Abstract

Spectrophotometric measurements at 0·54 *and* 0·69 μ *of the radiation from the 'inside' and the 'outside' of the helical filament of a* 12 V, 48 W *car bulb run at* 4 A *show that, owing to multiple reflection within the helix, the luminance of the inside is about* 70% *greater, and its colour-temperature* 50 $\pm$ 1 *deg* K *lower, than that of the outside. The colour-temperature of the integrated light from a helical filament depends critically on the angular position of the receiver relative to the source. Irregular variations of about* 10 *deg* K *have been measured within* $\pm 10°$ *of the equatorial plane of the filament mentioned, and similar variations for a* 100 V, 500 W *uniplanar projection lamp. Caution is indicated in using such lamps as standards of colour-temperature.*

INSPECTION of a magnified image of an incandescent coiled tungsten filament shows that the luminance of the 'inside' surface of the helix is much greater than that of the 'outside'. This is due to multiple reflections of radiation within the helix (Langmuir 1915, 1916, Shackelford 1915, 1916, 1922, Coblentz 1918–19, Holst *et al.* 1928). The inside of the helix is, in fact, a much better approximation to a full radiator than is the outside (or the surface of a straight filament) (Moon 1936).

Since the emissivity of tungsten is not quite non-selective, decreasing from about 0·46 at 0·4μ to 0·42 at 0·7 μ, at 2800° K (De Vos 1953, Larrabee 1957), the reflectivity must show a corresponding increase with wavelength, and multiple reflections must cause a reduction in colour-temperature of the radiation emerging from inside the helix. Apart from its academic interest, knowledge of the magnitude of this effect bears on the suitability or otherwise of coiled-filament lamps as standards of spectral distribution (standards of colour-temperature; standard illuminants).

The problem is hardly amenable to calculation (Helwig 1938, 1953), but, with the photoelectric spectrophotometric technique recently developed in connection with the N.P.L. scale of colour-temperature (Gibbs, to be published) direct measurement is possible, and has been made. Being capable of detecting, with certainty, differences of 1 deg K, the method is particularly suitable for this type of investigation.

A single-coil car-headlight bulb rated at 12 V, 48 W was mounted with its filament horizontal, and operated at a current of 4 A maintained constant to 1 in 10^4 by manual control monitored by a potentiometer. A 20 × magnified image of the filament was projected on to the 1 cm × 1 mm entrance slit of a Müller–Hilger Uvisir double monochromator. By means of a horizontal slow-motion screw on the source mounting, it could be arranged that the entrance slit selected radiation exclusively from the inside, or exclusively from the outside, of the helix, as shown in Fig. 1. The exit slit was imaged on to the cathode of a Rank-Cintel VB59 photocell, a type designed in collaboration with the N.P.L. for freedom from bulb-charge effects (Preston and Gordon-Smith 1955). The linearity of the particular cell used had been verified to better than 1 in 10^3. The response of the photocell was measured by means of an electrometer-valve bridge with a coupling resistor of 10^{10} Ω.

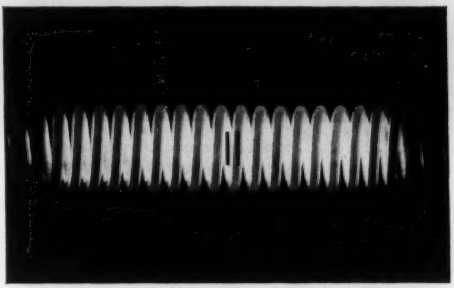
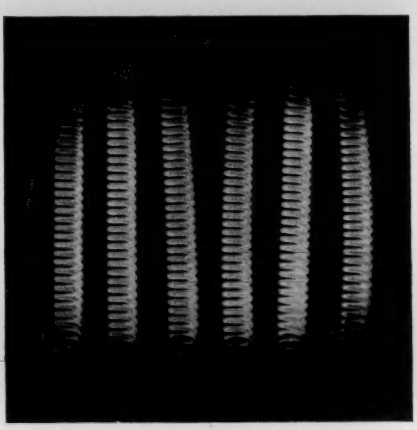


Fig. 1. Projected image of helical filament, showing the two rectangular areas sampled by the entrance slit.

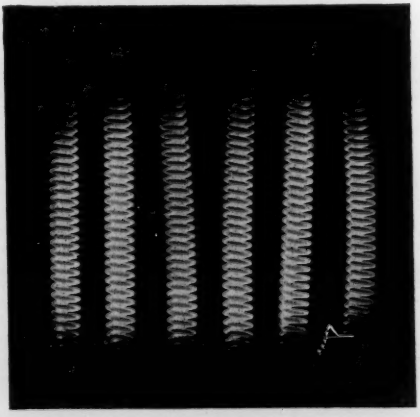
The colour-temperatures of the 'inside' and 'outside' samples of radiation were compared by reference to measurements of the ratio, for each of the two samples, of the spectrophotometric responses at nominal wavelengths of 0·54 μ and 0·69 μ. This ratio was interpreted in terms of colour-temperature by means of similar measurements on the integrated light from a standard lamp previously calibrated against the N.P.L. scale of colour-temperature (Gibbs 1961). In the spectrophotometry of selected areas of the car-bulb filament, and in that of the standard lamp, the angular distribution of radiation within the aperture of the monochromator would have been considerably different. The selectivity of the monochromator depends considerably on the angular distribution. A diffusing screen was therefore placed in contact with the entrance slit. It had previously been shown that, with this arrangement, symmetrical sources successively placed on the axis of the monochromator may be compared for colour-temperature with an error not exceeding 1 deg K, irrespective of their pre-slit angular distributions in the aperture.

In this way it was found that the colour-temperature of the 'outside' of the helix was 2940° K and that of the 'inside' 2890° K. It was inferred from the spectrophotometric data that the luminance of the 'inside' was about 1·7 times that of the 'outside'.

If the axis of symmetry of the helix be taken to define spherical polar coordinates, it is clear that the colour-temperature of the integrated light from the filament will be scarcely dependent on azimuth ϕ but strongly dependent on polar angle θ, where (θ, ϕ) denotes the angular position of



(*a*)



(*b*)

Fig. 2. Views of uniplanar projector-lamp filament with (*a*) top tilted 6° away from observer ($\theta = 96°$),(*b*) top tilted 6° towards observer ($\theta = 84°$).

a distant receiving element, for, as θ varies, more or less of the inside of each turn is occulted by the adjacent turn of the helix (Fig. 2).

The dependence of the colour-temperature of the filament on θ within 10° of the equatorial plane was investigated experimentally both for the single-coil car bulb and for a uniplanar 100 v 500 w projector lamp of a type often used as a standard of colour-temperature. The latter had six parallel helices. Errors of coplanarity of the axes of these six helices might be expected to cause an averaging of the separate effects and a consequent reduction in the θ-dependence of the integrated colour-temperature. In this lamp, however,

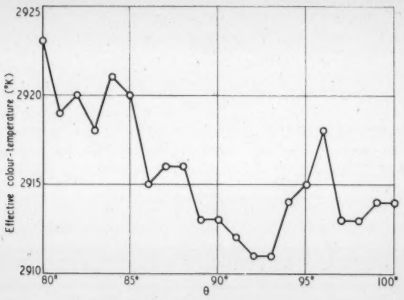


Fig. 3. Angular distribution of colour-temperature of a 12 v 48 w single-coil lamp.

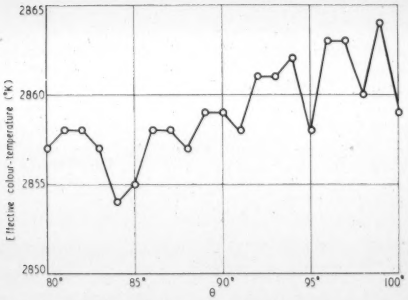


Fig. 4. Angular distribution of colour-temperature of a 100 v 500 w uniplanar projector lamp.

inspection showed that the six helices were closely coplanar, and little averaging seems to have occurred. The results, presented graphically in Figs 3 and 4, refer to a 2mm square receiving area at about 15 cm from the source.

Acknowledgment

The work described above has been carried out as part of the research programme of the National Physical Laboratory, and this paper is published by permission of the Director of the Laboratory.

References

COBLENTZ, W. W., 1918–19, *Bull. Nat. Bur. Stand., Wash.*, **14**, 118.

DE VOS, J. C., 1953, Thesis, Amsterdam, 54.

HELWIG, H.-J., 1938, *Das Licht*, **8**, 147.

—— 1953, *Technisch-wissenschaftliche Abhandlungen der Osram-Gesellschaft*, **6**, 107.

HOLST, G., LAX, E., OOSTERHUIS, E., and PIRANI, M., 1928, *Z. Techn. Phys.*, **9**, 192.

LANGMUIR, I., 1915, *Phys. Rev.*, **6**, 149.

—— 1916, *Phys. Rev.*, **7**, 326.

LARRABEE, R. D., 1957, *Research Laboratory of Electronics, M.I.T. Technical Report*, **328**, 39.

MOON, P., 1936, *The Scientific Basis of Illuminating Engineering* (New York: McGraw-Hill), 110.

PRESTON, J. S., and GORDON-SMITH, G. W., 1955, *Brit. J. Appl. Phys.*, **6**, 329.

SHACKELFORD, B. E., 1915, *J. Franklin Inst.*, **180**, 619.

—— 1916, *Phys. Rev.*, **8**, 470.

—— 1922, *Abstr. Bull. Nela Res. Lab.*, **1**, 329.

Notes and comments

Lay-out of journal

Readers will notice that the lay-out of the journal has been slightly altered in that each paper now begins on a fresh page. The Committee has decided to start this new system at once in order to increase the speed of publication and of reproduction of reprints and also to improve the look of the journal. As a result of this change a few book reviews may be published each month. We hope this will please those readers who prefer to see a few book reviews at a time rather than a large number in a special New Books section.

Noise in Electronic Devices

In October 1959 the Electronics Group held a short conference on 'Noise in Electronic Devices'. The seven papers there presented have now been published by Chapman and Hall Ltd., on behalf of the Institute and Society. The retail price is 35*s*. Members, Students and subscribers may obtain one copy each, for their personal use, at the reduced price of 24*s*. Their orders, with remittances, should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1.

Waverley Gold Medal Essay Competition 1961

Research is this year sponsoring The Waverley Gold Medal Essay Competition for the ninth year in succession. The Competition is designed to encourage the scientist in the laboratory and the engineer in the production plant to express his views and translate his work into an essay that will be readily understood by other scientists, directors of industrial firms and others interested in science and technology.

The Waverley Gold Medal, named after and bearing the coat of arms of the late Lord Waverley, together with £100 will be awarded for the best essay of about 3 000 words describing a new project or practical development in pure or applied science, giving an outline of the scientific background, the experimental basis and the potential or actual application of the idea to industry or their importance to society. The essays will be judged for technical content by specialists in the subject, for clarity of presentation and for style. The entry should be written as an essay which should interest a well-informed layman. In assessing these essays particular attention will therefore be paid (*a*) to the logical presentation, (*b*) to the style, and (*c*) to adherence to these terms of reference.

A second prize of £50 will be awarded and also a special prize of £50 for the best entry from a competitor under the age of thirty on 31st July 1961. If the first prize is awarded to a competitor under the age of thirty, the special prize will go to the next best entry. Entry Forms can be obtained from the Editor of *Research*, 88 Kingsway, W.C.2. *The last date of entry is* 31*st July*, 1961.

Modern developments in heat transfer

The University of Minnesota announces a second summer course in Modern developments in heat transfer, to be held at the University of Minnesota, Duluth, from 14th to 22nd August 1961. The first special course was presented by the University in 1958. The object of this second course is to provide a clear, basic exposition of the fundamental physical mechanisms underlying the various heat and mass transfer processes as some of the recent developments in the field are explored.

A preliminary outline of the programme is as follows: General review and exposition of the current status and problems in heat and mass transfer; Inter-relations between heat, momentum and mass transfer; Radiation; Plasma physics; Heat transfer with phase change (boiling and condensation); Mass transfer cooling; High-speed heat transfer; Properties for heat transfer calculations; Numerical methods in heat transfer.

The course will be directed by Dr. E. R. G. Eckert, Professor of Mechanical Engineering, Director of the Heat Transfer Laboratory, University of Minnesota. Further information may be obtained from the Director, Center for Continuation Study, University of Minnesota, Minneapolis 14.

Symposium on Inorganic Polymers

A Symposium on Inorganic Polymers is being arranged by The Chemical Society, London, in association with the University of Nottingham, and will be held from 18th–21st July 1961 at Nottingham. The main lectures will subsequently be published by the Chemical Society in the series of Special Publications. Applications to attend the symposium should be made to the Secretary of the Chemical Society, Burlington House, London, W.1, before 10th June 1961.

Post-Graduate Course in Welding Technology

The Institute of Welding announces the organization of a new one-year Post-Graduate Course in Welding Technology, to be held at the College of Aeronautics, Cranfield, starting in the Autumn Term 1961.

The course is a compound of mechanical and civil engineering and metallurgy and is organized with the object of producing a specialist in welding science and technology.

Forms of application may be obtained from The Warden, The College of Aeronautics, Cranfield, Bletchley, Bucks.

Applied Optics

Applied Optics, published by the Optical Society of America, will be issued bi-monthly from January 1962. It will contain a leader review paper on one of the major aspects of optics or its applications—this review will be about 10,000 words and will set the main subject of each issue; in addition contributed technical papers (about ten in each issue), shop and technical notes, letters to the editor, book reviews and news columns will be included.

Each issue will concentrate primarily on one subject; the first eight issues will probably be on Optical pumping, Space optics, Foreign optics, Optical engineering, Infra-red, Information theory, Astronomy and Instrumentation.

All orders for a three-year period, placed prior to publication, will be entered at a special rate of $15.00 for Optical Society of America Members and $25.00 for Non-Members for the three-year period. Normal yearly subscription rates after that will be $6.00 for O.S.A. Members and $10.00 for Non-Members. For further information apply to the Optical Society of America, 1155 Sixteenth Street, N.W., Washington 6, D.C., U.S.A.

Mining and Chemical Products Ltd.

Mining and Chemical Products Ltd. announce the publication of four technical booklets on the following subjects: Semiconductors; Thermoelectric materials; Bismuth telluride semiconductors; Galvanomagnetic effects and materials. Copies are available on request to Mining and Chemical Products Ltd., 86, Strand, London, W.C.2.

Soviet Physics-Doklady

The American Institute of Physics announce that their journal *Soviet Physics-Doklady*, now published every other month, will appear monthly, starting with the July 1961 issue.

The journal is a translation of the physics section of *Doklady Akademii Nauk SSSR*—the *Proceedings of the Academy of Sciences, USSR*. Editorial policy for this periodical emphasizes original research, breadth of coverage, brevity of contributions and speed of publication. Enquiries regarding the journal should be addressed to the American Institute of Physics, 335 East 45 Street, New York 17.

175 years Verlag Friedr. Vieweg & Sohn, Braunschweig

On 1st April 1786, Friedrich Vieweg established the publishing house which still bears his name today. During its latter years the firm has mainly dealt with the publication of books and periodicals in the field of mathematics, natural sciences and engineering, and with mathematical–scientific textbooks and special books. On the occasion of their 175th anniversary the publishing house is giving prizes for important papers in the field of mathematics, physics and chemistry, suitable for publication in book form. The publishers will be glad to send on request details concerning the conditions for the competition.

National Committee for Chemistry (Canada)

The XVIIIth International Congress of Pure and Applied Chemistry will be held in Montreal from 6th–12th August 1961, in connection with the XXIst Conference of the International Union of Pure and Applied Chemistry (2nd–5th August). The programme will consist of five plenary lectures and approximately fifty invited sectional lectures; the main divisions of the programme will be physical chemistry, applied chemistry, analytical chemistry and organic chemistry symposium. Further information concerning the Congress may be obtained from the National Research Council, Ottawa.

Journal of Scientific Instruments

Contents of the May issue

INVITED ARTICLE

The Annual Exhibition of The Institute of Physics and The Physical Society—London, 1961. By A. M. Taylor.

PAPERS

Stabilization of thermometers of borosilicate glass for use at high temperatures. By J. A. Hall and V. M. Leaver.
High-sensitivity recording optical density meter. By K. E. Collins and D. J. Steele.
Oxide resistor furnace for high-temperature operation. By E. Rothwell.
Method of modulating a high-intensity searchlight. By J. R. Bibby.
Indium resistance thermometers. By B. Yates and C. H. Panter.
Vacuum radiation furnace with precise control of temperature gradients for crystal growth by sublimation. By A. C. Prior.
Photoelectric optical pyrometer. By J. Middlehurst and T. P. Jones.
Design for a three-circle x-ray goniometer. By R. W. H. Small and S. Travers.
Sensitive method for measuring small rotations of a distant object. By R. J. King, S. P. Middleton and I. M. G. Thompson.
Automatic compensation for thermal e.m.fs and galvanometer zero drift in a feedback fluxmeter. By T. M. Palmer.
Nuclear resonance magnetic field stabilizer. By R. G. Eades, G. J. Jenks and A. Bradbury.
A sensitive electrostatic meter with a balanced reed. By Y. L. Yousef and S. A. Zaher.
A plastic capsule technique for the combustion calorimetry of volatile or chemically reactive compounds: the heat of combustion of polythene. By H. Mackle and R. G. Mayrick.

LABORATORY AND WORKSHOP NOTES

A simple and cheap thermal conductivity apparatus. By D. W. Stops.
Electropolishing technique for the preparation of zirconium and Zircaloy-2 specimens for transmission electron microscopy. By J. L. Whitton.
A method for sealing a watch jewel into hard glass. By G. P. Burn and P. J. Tosh.

CORRESPONDENCE

A.C. type ear oximeter. By T. Asada, K. Sugihara, T. Nakatsuka and T. Takeda.

NOTES AND NEWS

Notes and comments New books New instruments, materials and tools

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.
EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.
REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.
ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611–2.)
SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12*s.* 6*d.* each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.
CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

PRINTED IN GREAT BRITAIN BY UNWIN BROTHERS LIMITED, WOKING AND LONDON

Summarized Proceedings of a Conference on Physics of Polymers—Bristol, January 1961

Abstract

A conference on the physics of polymers arranged by The Institute of Physics and The Physical Society was held at The H.H. Wills Physics Laboratory of the University of Bristol on 10th–12th January 1961. The papers presented in the five sessions, devoted to Chain Statistics and Solution Properties, Molecular Motions, Crystallinity (two) and Irradiation Effects, together with an Evening Lecture by Sir Gordon Sutherland, F.R.S., are summarized in this report.

Introduction

DELEGATES were welcomed to the conference by Professor M. H. L. Pryce (H. O. Wills Professor of Physics, University of Bristol) who expressed the hope that the meeting would be both timely and useful for co-ordinating the diverse elements of this rapidly expanding field.

Professor R. Ullman (Brooklyn Polytechnic, New York) delivered the introductory lecture of H. Mark and R. Ullman. After apologizing for Professor Mark's absence, Professor Ullman began a survey of 'Expanding Areas in the Physics of High Polymers' emphasizing particularly those topics less well represented on the programme. At first he recalled the impetus polymer science had received from the synthesis of stereo-specific polymers and discussed ways of determining tacticity. An average effect was shown by certain gross properties such as melting point and dielectric constant. In special cases chemical reactivity, e.g. the rate of ester hydrolysis was altered. For local tacticity, high resolution nuclear magnetic resonance had been used by Bovey and Tiers on polymethylmethacrylate and on the related polyacid by O'Neill and Loebl. This could, in principle, be very useful for studying the configurations around successive carbon atoms in the polymer chain. Regarding new polymers, Professor Ullman singled out those which are highly conjugated. These systems are heat resistant, conduct electricity and show anomalous association in solution, though they are often insoluble and difficult to study. The high temperature stability of cyclized dehydrogenated polyacrylonitrile, of the anomalous viscosity of polyphenylazo compounds prepared by A. Berlin and of the electrical resistance of the copper polyphthalocyanine first synthesized by C. S. Marvel were quoted. The use of light and x-ray scattering in investigating solid structure was emphasized, mentioning the Guinier, Hosemann and Kratky–Porod interpretations of data. Some of the difficulties of calculating surface area from scattering data were examined by introducing a high angle cut-off (to simulate experimental conditions) on a simple model and the importance of the high angle data stressed. Finally attention was drawn to the examination of bulk properties in the light of new morphological discoveries. As an example, some photomicrographs of fracture in spherulitic films were shown where differences in morphology seemed to cause different modes of fracture.

Chain Statistics and Solution Properties

The first session of the conference 'Chain Statistics and Solution Properties' began with a general paper by Professor G. Gee (University of Manchester) on 'The Thermodynamics of Polymer Solutions' in which he confined himself to amorphous, non-polar polymers. He pointed out that these differed from liquid mixtures only by reason of the high molecular weight of one component. This difference was reflected in the entropy of mixing, and the ideal solution law had to be replaced by one appropriate to long flexible molecules. A very simple statistical calculation gave a law which reproduced all the main experimental features but which required further quantitative study, especially in three regions: (1) dilute solutions, where neither theory nor experiment was very satisfactory, (2) the absorption of moderately polar liquids which appeared to involve highly non-random mixing, and (3) the mixing of two polymeric fluids, where the limited experimental evidence available indicated unexpectedly low entropies of mixing.

A second general paper 'Chain Statistics—Introductory Lecture', was given by Professor C. Domb (King's College, University of London) surveying the field of the four papers following, which all adopted one particular approach to the theory of polymer configurations using a lattice model. Exact data for polymers of relatively few links were used and once initial small-number irregularities were damped out asymptotic formulae could be fitted. The method worked very well for the susceptibility of the two-dimensional Ising model where predictions could be checked against a known exact solution. Results obtained for the number of different polymer configurations and the mean square end to end chain length $\langle R_n^2 \rangle$ for n links, taking account of excluded volume, compared favourably with those of Wall and collaborators using Monte Carlo methods. (In two dimensions $\langle R_n^2 \rangle \simeq An^{1.48}$ for n steps and $\langle R_n^2 \rangle \simeq A'n^{1.20}$ in three dimensions.) Short range intermolecular forces could also be taken into account. The recent programming of a computer to enumerate configurations would, it was hoped, extend the method and enable long range electrostatic forces to be considered.

Dr. M. E. Fisher (King's College, University of London) considered the 'Statistics of Polymer Molecules in Dilute Solution' when the molecules are well separated. The principal feature of this work in which exact numerical data for short chains were extrapolated was the introduction of a Boltzmann factor $\eta = \exp(-V_0/kT)$ for every nearest neighbour contact of the polymer with itself, thus taking account of polymer–solvent interaction. The excluded volume effect was also included. $\langle R_n^2 \rangle$ was found to vary as An^γ where γ depended on η, i.e. on solvent and temperature. The existence of a Flory Θ point (at which $\gamma = 1$) was clearly indicated at $kT_\Theta/V_0 \simeq 0{\cdot}26$ in three dimensions and $\simeq 0{\cdot}14$ in two dimensions. The free energy for a chain of n links was found to vary as $-F_n/kT \simeq n \log \mu(\eta) + \alpha(\eta) \log n$ (the latter term describing the excluded volume) and could be estimated to within 1%. Accurate calculation of the entropy and internal energy had revealed serious errors in Orr's early estimates.

Mr. B. J. Hiley (King's College, University of London) gave a paper on the 'Configuration of a Polymer Chain with Excluded Volume'. He discussed a method for the exact

calculation of $\langle R_n^2 \rangle$ in dilute solution with excluded volume considered. This varied as An^γ and the fractional variance of the distribution of $\langle R_n^2 \rangle$ was found to be appreciably smaller than the corresponding Gaussian value.

Following, Dr. M. F. Sykes (King's College, University of London) spoke on 'The Probability of Initial Ring Closure in the Lattice Model of a Macromolecule'. He suggested that the Monte Carlo investigations of Wall and his co-workers, of ring closures in the lattice model of a macromolecule could be usefully supplemented by a study of the known exact probabilities of ring closure for short non-self-intersecting random walks. The behaviour of these was found to be sufficiently regular to enable extrapolations to be made and the results obtained were close to those of Wall. It was hoped to develop this method as more data became available from machine calculations.

'The Use of Electronic Computers for the Exact Enumeration of Self-Avoiding Walks' was the subject of Dr. J. L. Martin's (N.P.L., Teddington) paper. He stressed the need for subtlety to reduce the very large number of configurations to be counted. Suitable ways were (1) a flexible programme to handle different lattices without rewriting, (2) using auxiliary configurations such as 'tadpoles' which could save time by factors of up to 150, (3) arranging walks in a simple natural order and (4) the use of symmetry factors (sometimes as large as 192). The application of one programme was described which had given results as far as twelve steps on the quadratic, triangular and honeycomb lattices.

Opening the discussion, Professor G. S. Rushbrooke (King's College, Newcastle-upon-Tyne) mentioned results comparable with those of Domb and Sykes obtained at Newcastle. Professor F. C. Frank (University of Bristol) drew attention to the fact that the solvent contribution to the entropy of mixing in dilute solution was always greater in practice than the polymer contribution since there was not enough matter in the universe to dilute one high polymer molecule sufficiently to make the contributions equal. Professor Ullman raised the question of the validity of the extrapolation on the lattice model because there were early small closure problems whereas excluded volume, one felt, was related to long closures. Replying, Professor Domb mentioned the accuracy of the method in two Ising models where the answers were known and in agreement with Wall's Monte Carlo calculations on an 800 step walk. Dr. Fisher showed how a linear extrapolation could be extended farther and farther as more and more restrictions were removed in the calculations.

Molecular Motions

The second session was devoted to 'Molecular Motions', beginning with a paper by Professor A. Peterlin (Technische Hochschule, Munich) entitled 'Non-Newtonian Intrinsic Viscosity of Polymer Solutions'. He stated that completely soft molecules expanded in laminar flow. As a consequence of non-uniform expansion in which the close-by segments increased their average extension much less than the ends of the molecules, the hydrodynamic interaction first rose yielding a net decrease of intrinsic viscosity. At very high gradient, however, this effect reversed, so that intrinsic viscosity increased with the gradient. Such a dilatancy was found in extremely viscous polymer solutions ($= 10^4$ P). The finite length of actual macromolecules and their resistivity against rapid change of shape, however, suppressed to a large extent this rise of intrinsic viscosity. In later discussion, Professor Peterlin emphasized that he had treated one limiting case of viscosity. When molecules were not completely soft, inner viscosity due to potential barriers (which was the other limiting case) had also to be considered.

'Stress–Birefringence Relations for Flowing, Concentrated Polymer Solutions' was the title of the paper given by Dr. A. S. Lodge (British Rayon Research Association, Manchester). He described the first results using three new pieces of apparatus to measure the stress ellipsoid and a section of the refractive index ellipsoid in the steady shear flow of concentrated polymer solutions. The three methods of measuring the stress ellipsoid, using pressure distributions in cone-and-plate and parallel plate systems and the total thrust in a cone-and-plate system, gave discordant results for the orientation in a solution of polyisobutylene in decalin. The values obtained for this orientation were all different from the value for the orientation of the refractive index ellipsoid obtained from flow birefringence measurements made in a concentric cylinder apparatus. In reply to questions, Dr. Lodge said that attempts to extend rubber elasticity theory to concentrated polymer solutions were being made in an effort to provide a molecular basis for the experimental results.

Professor R. M. Barrer (Imperial College, University of London) followed with a paper 'Some Aspects of Diffusion of Penetrants in Polymers'. This phenomenon was, he said, a valuable means of investigating the degree of crystallinity, amount of cross-linking or of fillers and of the extent of chain mobility in polymers. Equally interesting were the effects of size and shape of the penetrant molecule and its polarity or cohesive energy density on the diffusion process. In a two-component diffusion system, five diffusion coefficients might be measured. The relations between these were considered and four categories of polymer–penetrant system then described in terms of the relevant diffusion equation, the behaviour of the measured diffusion coefficients and the nature of the equilibrium sorption isotherms. The penetrant diffusion coefficients were usually the important ones and could be strongly concentration dependent. In polymers below their glass transition temperature so-called anomalous diffusion appeared in which viscoelastic relaxation of swollen polymer occurred on a time scale comparable with that of the experiment. This gave time-dependent diffusion coefficients and could lead to two-stage sorption processes the first of which was governed by the rate of diffusion coupled with elastic swelling and the second stage, in which there might be practically no concentration gradients in the polymer, by the tempo of viscous relaxation coupled with further swelling and solvent uptake.

A paper 'Nuclear Magnetic Resonance Studies of Disordered Regions in Solid Polymers' was given by Dr. W. P. Slichter (Bell Telephone Laboratories, Murray Hill, N.J.). He had examined segmental motions in solid polyethylene and found that for melt-grown polymers the mobility in the disordered regions decreased with increasing crystallinity. In solution-grown crystals, heating produced chain mobility which varied reversibly with temperature below about 100° C. With higher temperatures, the mobility persisted after cooling. The abundance of mobility could be made to exceed that found in melt-grown samples and was insensitive to annealing. Studies of the nuclear spin–lattice relaxation time showed marked differences between all these materials with respect to the spectra of Brownian motions in the disordered regions.

In the discussion, Dr. Slichter stated that crystalline–amorphous ratios could be determined if the absorption frequencies of the two regions were separated; they were in polyethylene but not in Nylon 66. The question was raised of differences between dilute solution and melt crystallized specimens, which this work had shown not to be identical.

Dr. L. Mandelkern (National Bureau of Standards, Washington, D.C.) referred to recent results of Chiang and Flory who had found differences in melting temperatures of 12 to 15° C between the two types of specimen. At this point fuller discussion had to be curtailed.

The next paper 'Dielectric and Mechanical Properties of Polyoxymethylene' by B. E. Read and G. Williams (N.P.L., Teddington) was read by Dr. Williams. They had measured the dielectric absorption between 100 and 9×10^9 c/s from −80 to 160° C finding a single broad absorption process for which relaxation times and activation energies had been evaluated. With a polymer swollen in dioxane the relaxation was found to move to higher frequencies at constant temperature and lower temperatures at constant frequency. Detailed results indicated that the relaxation process arose from the amorphous regions and was compatible with a predominantly helical structure in those regions, with an increased probability of *trans* configurations at higher temperatures. The dynamic shear modulus and loss factor were determined from −190 to 180° C over a range of 0·05 to 1 c/s and two loss maxima found, at −77 and 87° C (0·1 c/s), the lower temperature maximum correlating with the dielectric absorption. Swelling in dioxane shifted the low temperature loss maximum to lower temperatures and broadened the dispersion. An increase in density (crystallinity) had the reverse effect and also lowered the modulus increment for the low temperature process. Both transitions were suggested to be associated with the amorphous regions and their temperatures with the free volume present. Thus it was supposed that the amorphous regions were in compression from the surrounding crystallites.

Dr. A. Elliott (Courtaulds Ltd., Maidenhead) read a paper 'The Omega Structure of Synthetic Polypeptides' with particular reference to poly-β-benzyl-L-sipartate. The polymer could be cast from chloroform solution in an oriented poorly crystalline form. X-ray diffraction and polarized infra-red spectra showed that the polypeptide chains were in the α-helix form. On heating and subsequent cooling, evidence of a different form of helix was found. This helix had a four-fold screw axis and packed tetragonally in what was called the ω-form. A model structure could be built which satisfied the requirements of the observations reasonably well, but a completely satisfactory solution had not yet been found. A two-dimensional Fourier synthesis had been made, making certain assumptions, by optical diffraction methods. Further work was required, but the main features of the structure appeared to be established.

When discussion was opened, Dr. Elliott replied to a question on the molecular basis of the transition from the α-helix (3·6 residues per turn) to the ω-form in which there were four, that he supposed there was a greater interaction energy between the four-fold helices. To further questions, he said he had no evidence that the transition proceeded through a transitory liquid phase and that no change in optical rotation had been found for the transition.

Here the session ended and the conference adjourned to Churchill Hall where, after dinner, Sir Gordon Sutherland (Director of N.P.L., Teddington) gave an Evening Lecture, 'The Potentialities of Infra-Red Analysis in Polymer Research'. Professor Pryce took the chair. Sir Gordon illustrated his subject by a review of the contributions which infra-red methods had made to a representative selection of polymers. In polymers for which the repeating unit was small and the crystal structure was known, the infra-red spectrum of the fully crystalline polymer was no more complex than that of a simple polyatomic molecule. Strict selection rules operated and all the main features of the spectrum were now well understood. Although in the majority of polymers the analysis was, as yet, incomplete, the assignment of a number of key bands was sufficiently certain for them to be used to discriminate between alternative structures put forward on the basis of x-ray analysis (e.g. isotactic polystyrene and polyvinylidine chloride). Infra-red methods could also be very useful in detecting changes in polymers arising from irradiation and oxidation. The anomalies in the infra-red spectra of Type I diamonds were used to illustrate the power of infra-red analysis in revealing the presence of unexpected impurities. The differences between the infra-red spectra of the crystalline and amorphous forms of a polymer were stressed as was the need for more research on the reasons for these differences. Finally, a brief review was given of the ways in which infra-red analysis had contributed to our understanding of the structure of proteins and of desoxyribonucleic acid.

In the subsequent discussion, Sir Gordon mentioned that some work had been carried out at the N.P.L. on a sample of solution-grown polyethylene supplied by Keller in an attempt to see whether the folded chains could be detected by strains imposed on the methylene groups. None had been found and very careful dichroic measurements would be necessary for a thorough examination.

Crystallinity

The second day of the conference was devoted entirely to 'Crystallinity'. Papers read in the morning summarized the current experimental position and in the afternoon several theories of chain-folded growth were put forward. The first paper was a review 'Molecular Texture in Crystalline Polymers' read by Dr. C. W. Bunn (I.C.I., Welwyn Garden City). His lecture consisted of a general account of the evidence and a critical survey of suggested interpretations. Early x-ray diffraction evidence on material crystallized from the melt led to the 'bundle and tangle' molecular picture of the two phases (crystalline and amorphous), while optical and x-ray evidence demonstrated the presence of spherulites with tangentially oriented chain molecules. Electron micrographs suggested that each spherulite was built up by crystal growth from a single nucleus, the radial structure being the result of continued distorton and branching; the best evidence for continuity (rather than successive nucleation) was that when two different crystal forms grew in the same specimen, they were in separate spherulites, not mixed in the same spherulite.

Single crystals from solution (thin plates) often grew by the screw dislocation mechanism; electron microscope and x-ray evidence showed that the chains stood perpendicular to the layers and suggested strongly that the molecules are folded regularly at intervals of about 100 Å; there might be a little molecular tying between successive layers. Two theories of chain folding had been suggested—a kinetic theory in which the first nucleus was a folded molecule and the initial fold period was maintained in subsequent growth, and an equilibrium theory which suggested that in a crystal of long chain molecules, the thermal motions led to instability when straight sections of the chain exceeded a certain length. Fold structures opened a new chapter of polymer crystallography in which crystals could have symmetries different from that of the sub-cell, and could show twinned domains or even different phases in different regions of the same layer.

Polymers in use crystallized in bulk from the melt. Layers 100–200 Å thick were seen on the free surfaces; the orientation

of the chains in these layers had not been established by diffraction evidence, but spiral layers suggested that the chains stood perpendicular to the layers as in solution-grown crystals, and therefore chain folding was likely. For the interior, replicas of fracture surfaces showed little evidence of structure for most materials, but in polytetrafluoroethylene showed well-marked layers which, though much thicker than in other materials, may be evidence of chain folding if current evidence on molecular length (admittedly shaky) was accepted. It seemed unlikely that a completely folded structure without molecular ties between layers would be formed, if the melt had a random tangle structure. Moreover, molecular ties between layers or crystals appeared to be necessary to explain the toughness of polymer specimens, which contrasted sharply with the properties of short-chain substances having no molecular ties between layers, and which was the foundation of their practical usefulness. What was wanted was a method of estimating the proportions of chain folding and of 'bundle and tangle' structure in polymer specimens.

There followed three papers on polymer morphology; the first, entitled 'Observations on Polymer Single Crystals' by D. C. Bassett, A. Keller and S. Mitsuhashi (University of Bristol), was given by Dr. Keller. It had been demonstrated that in a given system the thickness of the crystals depended only on crystallization temperature. It increased or decreased apparently discontinuously within the same crystal as the temperature was raised or lowered respectively. This fact formed the experimental basis of the latest kinetic theory of polymer crystallization by Frank and Tosi.

Other observations showed that monolayer polyethylene crystals could be non-planar (Fig. 1). It was proposed that a particular non-flat-based 'dished' pyramid could account

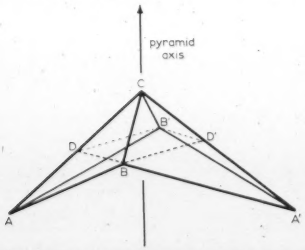


Fig. 1. Diagram of the non-flat-based hollow pyramidal model of a polyethylene crystal proposed by Bassett, Keller and Mitsuhashi. A,A', B,B' are the corners of the pyramid and C its apex. An imaginary plane through B,B' perpendicular to the pyramid axis would intersect the pyramid in BDB'D'.

for a variety of observations. This crystal shape could be correlated with the packing of the folds; it resulted from a particularly simple mode of staggering between adjacent folds. Pyramids like Fig. 1 and related configurations had been observed by dark field optical microscopy.

Thirdly, observations on multilayer crystals turning over in suspension had revealed a splaying of the lamellae (as in Fig. 2, showing a side view of a crystal with spiral terraces). This raised new questions regarding the connectedness of consecutive layers in multilayer crystals and pointed to a new possibility of generating spherulites (Fig. 3) via single crystals. A probable intermediate stage between the two, hedrites (Fig. 4), was discussed by Dr. P. H. Geil (E. I. DuPont de Nemours, Wilmington, Del.) in the next paper, 'Morphology of Polyoxymethylene'.

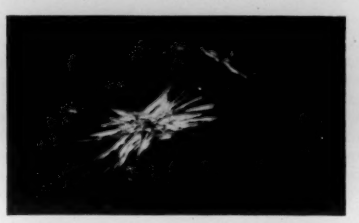


Fig. 2. Edge-on view of a polyethylene multilayer crystal in suspension showing the splaying of layers. Photomicrograph (×410). *After* Bassett, Keller and Mitsuhashi.

Fig. 3. Surface replica of polyethylene spherulites crystallized from the melt, showing lamellae. Electron micrograph (×7500). *After* Fischer.

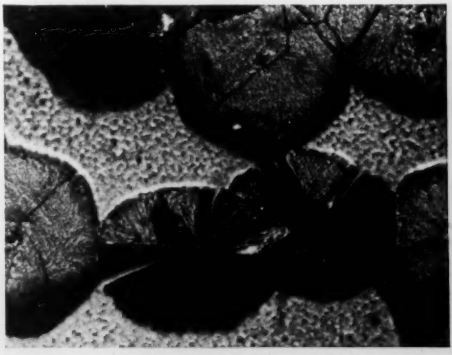


Fig. 4. Melt-grown polyoxymethylene hedrites and oval structures. *After* Geil (×500). Hedrites can be grown from both solution and melt.

This polymer crystallized from both melt and solution as chain-folded lamellae, the thickness depending on crystallization temperature and subsequent annealing. Simple crystals of a few lamellae could be grown from solution. The more complex hedrites (Fig. 4) with a large number of lamellae all with the same orientation and apparently originating at a screw dislocation had been grown both from

solution and from the melt. They gave a single crystal-like diffraction pattern. Melt-grown spherulites grew similarly but more rapidly and irregularly. First a single lamella grew in which eventually a screw dislocation formed, followed by the growth of secondary lamellae initially with the same orientation (which was maintained during hedrite growth). Irregular twisting and branching of the lamellae leaving the nucleus created tertiary lamellae near the periphery of the growing spherulite whereas secondary lamellae originated at the nucleus.

Finally, Dr. E. W. Fischer (University of Mainz) reported 'On experiments about Chain Folding and Large Periods in Drawn Fibres' which had been done in conjunction with G. F. Schmidt. They had examined the dependence of the fold period of polyethylene crystals on crystallization temperature and solvent and of the long period of drawn polyethylene films on temperature and time. Low-angle x-ray results on sedimented films of polyethylene crystals crystallized between 50° and 110° c had revealed that to a good approximation the fold period was independent of solvent, but its decrease with increasing supercooling was not so great as the kinetic theories of Lauritzen–Hoffman and F. P. Price had predicted. The long period in drawn polyethylene films increased similarly to that of annealed crystals. The time dependence showed that after a few minutes a value was reached which changed only slightly after days of further heating. These results, it was suggested, favoured a thermodynamic explanation of crystallite size along the chains.

The kinetics of crystallization was the concern of the next two papers, the first by Dr. M. Gordon (Imperial College, University of London) 'On Crystallization of Polyethylene'. Dilatometric studies had been made in collaboration with W. Banks, J. N. Hay and R. J. Roe on crystallization and melting kinetics for two sharp Marlex 50 fractions in the range 124 to 130° c. The Avrami exponent n as a function of crystallization progress was derived by an objective method to at least $\pm 0{\cdot}1$; n rose initially by one to two units, probably due to changes in the nuclear shape, and then levelled off to a fractional value constant over a decade of time. A high molecular weight fraction gave n decreasing smoothly with rising temperature and the Avrami rate constant fitted the equation

$$\ln z = A - \frac{nB}{T^2(T_m - T)^2}$$

where T_m was the *measured* melting point. The isothermal volume increase on melting was a rate process whose asymptotic behaviour at large times showed that during the post-Avrami (secondary) stage of isothermal crystallization, the crystallinity became progressively more stable (higher melting point).

'Crystallization of poly(decamethyleneterephthalate)' was the title of a paper by A. Sharples and F. L. Swinton (British Rayon Research Association, Manchester) read by Dr. Sharples. Above 123° c crystallization of the pure polymer conformed to Avrami kinetics with $n = 4$ indicating sporadic nucleation and spherulitic growth. At 122·5° c and below, the nature of the crystallization changed, the Avrami plot became curved and the density change on crystallization was 40% less at 122·5° c than it was at 123° c. It was not thought that a change in crystal form or contributions from pre-existing nuclei could explain the size of the effect. Most likely seemed to be a change in the spherulitic growth process though this could not be confirmed microscopically as the birefringence appearing on crystallization was not resolvable optically.

In discussion Dr. L. B. Morgan (I.C.I., Manchester) suggested that the deviation from n integral could be due to competing crystallization processes between the long methylene parts of the polymer chain and the terephthalate part of the structure. Competition would occur below the melting point of the polymethylene chain. Dr. Sharples thought this an interesting proposal, but found it difficult to see how a second crystalline phase could lead to a reduction in density. Dr. P. H. Geil mentioned work in his laboratories on similar polymers to those of Dr. Sharples in which two types of crystallization had been found.

The last paper of the morning 'Orientation in Nylon Spherulites: A Study by X-ray Diffraction' was read by Dr. J. Mann on behalf of himself and L. Roldan (British Rayon Research Association, Manchester). Microbeam studies on positively birefringent spherulites had shown an a axis radial orientation. The zigzag extinction pattern observed between crossed polars was compatible with this if account was taken of the imperfections in real spherulites. In disagreement with earlier workers, no orientation had been found in the x-ray patterns of negative spherulites. Non-birefringent spherulites were found to show random orientation. Spherulitic aggregates when grown in thin films ($\sim 10\ \mu$) gave a Maltese cross with arms at 45° to the vibration directions of the polars. An orientation was deduced from x-ray work consistent with this. In thicker films ($\sim 80\ \mu$) the optical properties were confused as the orientation varied from point to point in the aggregate, the spherulite radius having little significance for the orientation.

Dr. C. G. Cannon (British Nylon Spinners, Pontypool) reported in discussion work on Nylon 66 spherulites which seemed consistent with a progressive reduction in the number of orientational degrees of freedom. Thus non-birefringent spherulites grown at the higher temperatures had no restrictions, negative spherulites one, in that the c axes were tangential, and positive spherulites two with the a axis radial and c tangential.

The afternoon session was the occasion of some theoretical papers concerned with folded chain growth. Dr. J. I. Lauritzen (National Bureau of Standards, Washington, D.C.) contributed a paper 'On Theory of Chain Folding in Dilute Solution'. In sufficiently dilute solution entropy considerations predicted the preponderance of folded chain over bundle-like crystals from a linear homopolymer provided steric factors did not interfere. The step height of the primary nucleus would be $l_p^* = 4\,\sigma_e/\Delta f$, where σ_e was the surface free energy of the loop-containing surfaces and Δf the volume free energy change on crystallization. After growth, flat, plate-like crystals were predicted with the upper and lower surfaces containing the folds. Growth probably proceeded by the addition of monomolecular layers of chain-folded molecules on the lateral crystal surfaces, and certain aspects of this were considered. The distance between the folds was quite uniform and determined by the crystallization temperature.

Professor F. C. Frank followed with a paper 'Theory of Crystallization of Polymers' by F. C. Frank and M. P. Tosi (University of Bristol). This theory began with, but departed from, that of Lauritzen and Hoffman (previous paper) by showing that successive fold segments were not in fact of the same length but fluctuated considerably. An approximate calculation showed that within a specific range of cooling there was a characteristic segment length l^* such that after shorter segments than l^*, the next fluctuation was more likely to be a longer one and vice versa, so that l^* defined a stable mean strip width into which the molecule folded. The

next strip deposited on this one had a narrower stable width. With successive strips, their widths converged to a stable value l^{**} which depended on temperature in the observed way. For smaller supercoolings, the strip width converged to a value too narrow for any further growth to occur. At larger supercoolings, multiple nucleation of crystals within the same molecule was considered to change the mode of crystallization. Acceptable parameters gave satisfactory agreement with experiment.

Dr. J. D. Hoffman (National Bureau of Standards, Washington, D.C.) was unable to be present to read his paper 'Possibility of Chain Folding in the Bulk'. Dr. Lauritzen, however, gave a summary of Dr. Hoffman's conclusions. Very briefly, these were that the existing theories of chain folding in dilute solution could be extended to lamellar spherulitic crystallization from the melt for which there was an increasing amount of evidence.

Next, Dr. F. P. Price (General Electric, Schenectady, N.Y.) presented 'A Markoff Chain Model for Polymer Single Crystal Growth'. The growth of a polymer single crystal was taken as the successive coherent addition of layers of segments of molecules, each layer acting as substrate for the next. The substrate was assumed smooth and the distribution of segment lengths governed by equilibrium configurations. It was predicted that for increasing supercooling, the energy gained in crystallization from the volume term would tend to overwhelm that lost in creating new surface. This led to a temperature dependence of ultimate crystal thickness less than the $(\Delta T)^{-1}$ laws of previous theories. Agreement of theory with experiment was satisfactory. The theory also showed that the surfaces of the polyethylene crystal were remarkably smooth with an average roughness of only 1–2 Å.

Before the discussion, Professor A. Peterlin presented very briefly the equilibrium theory of Peterlin and Fischer. Their argument was that the incoherent part of the longitudinal and rotational fluctuations of polymer chains introduced in the free energy density a term increasing with the number N of elements in the straight portion of the chain. Together with the surface energy contribution the free energy density had a minimum yielding the number N^* of the most stable configuration. This number increased with increasing temperature and decreasing interchain forces in the crystal in good agreement with experimental data.

In discussion, Professor Frank explained that entropy contributions to crystallization did not dominate the resultant fold length because of the large energy of the folds. He went on to show how the fold energy could be estimated since at least three gauche bonds were required to reverse the polyethylene chain. The energy needed for this was known and together with further contributions to satisfy the sub-cell packing gave a fold surface energy of about 100 erg cm^{-2}.

Of the final three papers in the 'Crystallinity' section, the first was by Professor R. Hosemann (Fritz Haber Institute, Berlin-Dahlem) 'On Paracrystals'. Using a convolution polynomial of the three fundamental 'co-ordination statistics' the correlation function (expectation of the convolution square of the total distance statistics) could be calculated. This gave a generalized expression of the small range disorder in three-dimensional lattices, degenerating for point-like co-ordination statistics into the lattice function of a crystal. For special statistics in the one-dimensional case it gave the formulae of Zernike–Prins and J. J. Hermans and for three-dimensional unoriented lattices the equations of conventional theories of liquids. It was shown that the macrolattices of various fibroins (e.g. α and β keratin and collagen) were of a paracrystalline type and some of its 18 tensor parameters were calculated. Small-angle diagrams of variously treated polyethylene fibres, of aggregated polyethylene crystals and Fraunhofer patterns of paracrystalline models were analysed showing how quantitative information about the paracrystalline disorder in natural substances could be obtained.

Dr. C. Robinson (Courtaulds, Maidenhead) read a paper 'On Liquid–Crystalline Structures in Solutions of Polypeptides'. This type of phase appeared in several optically active polypeptides dissolved in organic solvents above a certain concentration. It was spontaneously birefringent, showed visible periodicities and very high optical rotatory power. The molecules in solution were rigid rods, being in an α-helix configuration. The actual structure in solution was derived from an arrangement of parallel molecules by superimposing on it an axis of torsion perpendicular to their long axes. The observed periodicity equalled half the pitch of torsion ranging between 100 μ and sub-microscopic dimensions depending on solvent, concentration and temperature. For suitably sized periodicities iridescent colours were reflected in accord with the Bragg equation. The observed optical rotation was consistent both with the proposed structure and the de Vries theory of cholesteric liquid crystals. There was also some evidence of similar structures in aqueous D.N.A. solutions.

Dr. L. Mandelkern contributed a paper 'On Contractile Processes in Fibrous Macromolecules' in which anisotropic dimensional changes were shown to accompany the crystal–liquid transformation in these substances. Contraction was observed on melting, elongation on crystallization due to the different conformations of the chains in the distinctly different states. The problem of carrying out these changes reversibly was solved in two ways, firstly by imposing a uniaxial tensile force and maintaining the two phases in equilibrium; secondly, cross-linking highly oriented fibres resulted in a substantial increase in the isotropic length so that after melting and recrystallization oriented fibres were again found. These were reversibly contractile without having to maintain an external force. These results could be applied to contractility in fibrous proteins.

In a written contribution to the discussion, Dr. J. Sikorski (University of Leeds) commented that the reversible supercontraction of wool fibres, of the order of 30%, could be realized (in certain metal amines) without any serious disturbance in their diffraction pattern. Under other conditions reversible contraction of only 2% was obtained (involving hydrogen bond breakdown) with a concurrent disappearance of the α-crystalline pattern. It was therefore not possible to generalize and suggest that loss of crystallinity was a necessary accompaniment of reversible supercontraction. One had to remember the chemical and morphological heterogeneity of wool fibres.

In a brief discussion, Dr. R. St. John Manley (McGill University, Montreal) showed some very interesting slides of crystals allegedly of cellulose triacetate and cellulose II. Cellulose triacetate when dissolved in nitromethane to more than 4% solutions precipitated as spherulites. Square, lamellar crystals could be obtained in more dilute solutions by slow cooling in the presence of a non-solvent. These crystals displayed many of the features familiar in polyethylene crystals. By deacetylating the triacetate crystals with alcoholic potash square crystals were retained with lamellae clearly visible at their edges. Aggregates of these gave an x-ray diffraction pattern corresponding to cellulose II.

Irradiation effects

The fifth and final session, Irradiation Effects, began with a paper 'Free Radical Kinetics in Irradiated Polymers' by A. Charlesby and M. G. Ormerod (Royal Military College of Science, Shrivenham), read by Mr. M. G Ormerod. Electron spin resonance spectroscopy had been used to study the free radicals formed in polyethylene irradiated *in vacuo* at 77° K. The predominant radical was the alkyl radical $-CH_2-\dot{C}H-CH_2-$. On warming to room temperature the radicals decayed. The main effects of radiation on polyethylene were the formation of cross-links and main chain unsaturation. Since main chain unsaturation was formed at 77° K the disappearance of the alkyl radical was associated with cross-linking. The G value for radical formation was 2·5 while that for cross-linking at 77° K was about 1. This supported the view that one cross-link was formed from two radicals. Experiments with two paraffins and an olefin suggested that the alkyl radical was mobile. They therefore postulated that a similar radical mobility occurred in polyethylene. Experiments were being performed to enable the degree of mobility to be estimated for polyethylenes with different physical properties.

The next paper, 'Transfer of Molecular Excitation Energy in Aromatic Polymers', was given by Dr. J. B. Birks (University of Manchester). According to his theory, the energy dissipated by ionizing radiation in an aromatic solvent was partitioned between excitation (10%) and ionization (5%) of π-electrons, and excitation (57%) and ionization (28%) of other electrons. The π-excitation was converted into the lowest solvent π-singlet state from which either emission, migration to another solvent molecule or transfer to a solute molecule occurred. Studies on polystyrene solution showed that solvent–solvent migration was inefficient compared with crystalline or liquid solvents and that solvent–solute transfer occurred radiatively below 10^{-4} M and both radiatively and non-radiatively above 10^{-4} M. Ion recombination after π-ionization yielded excited triplet states. The other primary processes which produced free radicals, electrons and excited molecules were responsible for the radiation-chemical effects.

Dr. W. L. M. McCubbin (King's College, Newcastle-upon-Tyne) gave a paper 'Electronic Processes in Paraffinic Hydrocarbons'. His reasons for rejecting earlier theories of cross-linking induced in paraffins by x-radiation were their inability to explain (*a*) the dependence of cross-linking efficiency on small percentages of impurity, and (*b*) the independence of this quantity on dose-rate or temperature. The suggestion made by Weiss was that a diffusing hole eventually came close enough to the hydrogen atom on a neighbouring chain thereby forming a cross-link with the elimination of hydrogen.

Two experiments were described which indicated that holes were, in fact, the majority carrier in irradiated hydrocarbons. It was shown to be possible to control the ratio of hole to electron components of the d.c. current by adding small amounts of electron-accepting and electron-donating impurities, i.e. benzo-quinone and dihydroanthracene respectively. Experiments to check whether the cross-linking efficiency was affected by these impurities were in progress.

A paper 'Radiation Effects in Polyethylene as a Function of Crystallization Conditions' by A. Keller and R. Salovey (Bell Telephone Laboratories, Murray Hill, N.J.) was read by Dr. Keller. Samples of Marlex 50 crystallized from dilute solution and from the melt were irradiated by high energy electrons and their solubilities compared. The single crystal specimens remained fully soluble after a 20 Mr dose which rendered the bulk $\frac{3}{4}$–$\frac{3}{4}$ insoluble. It was suggested that intramolecular cross-links occurring more in the single crystals were responsible for this behaviour. At high doses, the same solubility limit was approached by both specimens; this was ascribed to the increasing ineffectiveness of chain cutting in the single crystals.

Other Marlex 50 samples were solution-crystallized in the range 0·1 to 2% and given the same 20 Mr dose. Those crystallized at 70° C were largely soluble, but those crystallized at 85° C from concentrations more than $\frac{1}{2}$% became $\frac{3}{4}$ insoluble. Examination showed that crystallization was lamellar in both cases, but was much more compact at 85° C. However, after ultrasonic treatment which separated some of the lamellae the 85° C specimens stayed fully soluble. This meant that the radiation behaviour depended on the lamellar packing. Thus radiation effects could serve to distinguish modes of lamellar packing, which morphological factor might be responsible for the differences between single crystals and the bulk. The results further suggested that cross-links form largely at the lamellar interfaces.

Professor M. Dole (North Western University, Evanston, Ill.) spoke on 'The Radiation Chemistry of Polypropylene' about work in which he had been assisted by W. Schnabel, R. Keyser and B. Clegg. They had found that gel formation in irradiated atactic and isotatic polypropylene closely followed the Charlesby–Pinner formula. This formula was valid for a random molecular weight distribution and adherence to it meant that degradation had created such a distribution in the material. $G(S)$ and $G(X)$ values were calculated to be around 0·3 for a random initial distribution and around 0·1 for a 'pseudo-random' one. Black and Lyons had previously found $G(S) = 4{\cdot}95$ from calculations on the initial viscosity changes on irradiation. A new treatment by the authors taking various 'constants' in the defining equation as varying with dose had explained their own viscosity results and substantiated the low G values quoted.

Before reading the final paper of the conference, Professor A. Charlesby (Royal Military College of Science, Shrivenham) congratulated the sponsoring body and the organizers on the success of the conference. His paper 'Comparison of Ultraviolet and Gamma Radiation Effects in Polymers' had C. S. Grace and D. K. Thomas as co-authors. In an attempt to distinguish between ionization and excitation in solid polymers, results had been compared of irradiations with ionizing radiation and ultra-violet light which could cause only excitation. In polymethylmethacrylate, degradation occurred with both radiations being proportional to dose but independent of dose-rate. Electron spin resonance had shown that the radicals formed were the same in the two cases. In polyethylene, cross-linking, degradation and the formation of *trans*-unsaturation all occurred in approximately the same proportions in both cases. It was concluded that the main changes in the two polymers were due primarily to radicals formed by excitation, there being no evidence of any special contribution from the ions from gamma radiation. The energies absorbed to promote these reactions were, however, very different so that while the prime processes of energy absorption varied considerably, the course of the reaction soon ran along the same path. It might be, though, that other radiation reactions, e.g. conductivity, did show evidence for an ionic mechanism. After some discussion on this theme, the conference ended and the delegates dispersed.

H. H. Wills Physics Laboratory,
University of Bristol,
Royal Fort,
Bristol 8.

D. C. BASSETT

Assessment of the accuracy of 'flying-spot' scanning for the measurement of microscopic particles

by C. G. L. FURMIDGE, B.Sc., Ph.D., A.R.I.C., Department of Agriculture and Horticulture, University of Bristol, Long Ashton, Bristol

MS. received 29*th November* 1960, *in revised form* 6*th January* 1961

Abstract

A critical review has been made of the reliability and accuracy of 'flying-spot' scanning for the automatic measurement of microscopic particles. The instrument used was the 'Flying-spot particle resolver' manufactured by Messrs. Rank Cintel Ltd. Most of the results quoted were obtained using a reflected-light scanning technique for the measurement of spray droplet stains, but it is probable that the conclusions may be applied to the measurement of many other types of microscopic particles. The possible errors associated with automatic methods of measurement have been considered under the headings:
(a) errors inherent in the design of the instrument;
(b) errors due to faulty operation of the instrument;
(c) errors due to the nature of the sample being measured.

The accuracy and reliability of the automatic assessment of particle size has been compared with the visual assessment, using a standard optical microscope. It is concluded that the speed and reliability of automatic measurement are considerably greater than can be achieved by visual measurement. Basically the instrument is very accurate, but the accuracy with which a particular sample is assessed depends to a considerable extent on the nature of the sample. Inaccuracies are introduced by overcrowding or aggregation of the particles on the sample, by irregularly shaped particles and by the absence of well-defined, clear-cut edges to the individual particles on the sample. The importance of these possible inaccuracies is discussed.

1. Introduction

During recent years there has been considerable interest in the possibility of counting and sizing microscopic particles by rapid mechanical or electronic methods. The development of such methods was reviewed at the 1954 Institute of Physics conference on the physics of particle size analysis. The more successful methods that have come into prominence recently have depended on the photoelectric scanning of the particulate sample: either by using track-scanning (Hawksley 1954, Hawksley, Blackett, Meyer and Fitzsimmons 1954, Morgan and Meyer 1959) or spot-scanning systems (Taylor 1954, Causley and Young 1955).

Besides reducing the amount of mental strain that is involved in visual microscopic counting and sizing, automatic methods may minimize the personal variation that always occurs between different operators and may also prevent the loss of accuracy that is experienced when visual counting is carried out over a long period of time. The rapidity with which the counting is carried out automatically also means that more replicates may be measured, with a consequent reduction in sampling errors. On the other hand, automatic counting may suffer from certain disadvantages; the instrument may not be working accurately, it may be incorrectly adjusted and an artificial scanning technique is almost certainly not as discriminating as the human eye.

The performance of an instrument using a track-scanning method has been described by Morgan and Meyer (1959), and the work described in this paper is an attempt to evaluate the advantages and disadvantages of automatic counting and size determination using a commercial instrument that employs a spot-scanning technique. The instrument used was the 'Flying-spot particle resolver' manufactured by Messrs Rank Cintel Ltd. Since it was initially wished to use this instrument for the measurement of droplet size in agricultural sprays, most of the results were obtained from spray samples, but there seem to be no reasons why the conclusions should not be applicable to many other types of microscopic particles.

Despite the quantity of work carried out on the development of instruments capable of automatic counting and sizing, the available literature contains little information on the operation and accuracy of such methods from a practical viewpoint. This is probably due, at least in part, to the difficulty in finding suitable standards against which to compare an instrument's performance. Courshee (1954) discussed some of the practical difficulties associated with the operation of the chord intercept analyser developed by Trickett and Courshee (National Institute of Agricultural Engineering Report No. 27, 1953) and he gave some comparisons between visual and automatic counts. Courshee concludes that 'the experimental error associated with eye counting is rather larger than is often assumed in accepting this method as a comparison standard'. This conclusion is supported by the present work, but since visual counting and sizing is still most commonly used for the measurement of size distribution in microscopic particles it is useful for comparison with automatic counting and sizing, provided that it is not assumed that the accuracy of visual counting is absolute.

2. The 'Flying-spot particle resolver'

The operation of this instrument for visual presentation is shown diagrammatically in Fig. 1. A 700-line scanning raster (12·5 scans/s), produced on the face of the scanning tube, is passed into the optical system, which usually consists of a standard optical microscope. The image of the scan raster is then focused on the sample to be examined, using the normal microscope focusing device, the scanned area of the sample being determined by the power of the objective lens used. The amount of light passing through the sample

where transparent backgrounds are used, or reflected from the sample if it is opaque, varies according to the optical density and configuration of the objects on the sample, and these changes in light intensity are detected by a multiplier photocell in which they are converted into an electrical signal.

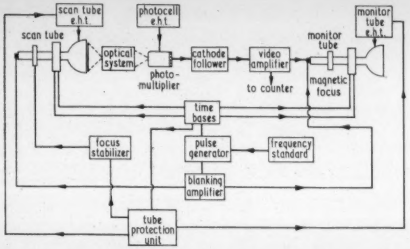


Fig. 1. Block diagram of circuits used for visual presentation.

This signal is fed to the video amplifier, the output of which modulates the monitor cathode-ray tube that operates in synchronism with the original scanning raster. Thus a magnified image of the sample is produced on the monitor screen, the degree of magnification being dependent on the lenses used in the optical system. Using transmitted light the magnification is variable up to 7000 diameters; with reflected light the magnification used is in the order of 20–50 diameters. The normal microscope controls for focusing and stage movement are used, and the contrast of the image is controlled by the video amplifier gain and by the brightness of the scanning and monitor tubes.

Automatic counting of all the particles within the scanned field is carried out by using the extra units shown in Fig. 2.

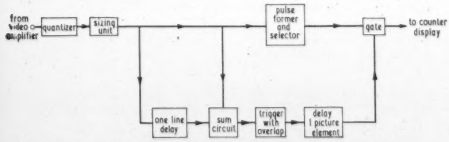


Fig. 2. Extra circuits required for operation on counting and sizing (see Fig. 1).

Considering the case of only one particle on the sample slide, when this is scanned by the flying spot an intercept pulse will be produced which passes to the quantizer. This unit will pass on the pulse only if it is greater than a predetermined voltage level (the clipping level); it thus prevents counts from spurious particles and random noise. From here the pulse passes to a magnetic memory system (the one-line delay), the sum circuit and the pulse-forming unit where it is suitably shaped. After shaping, the pulse passes through an electronic gate, which is normally open, and then operates a dekatron counter tube. Acceptance of a count pulse is shown on the monitor screen by a bright spot appearing on the particle. If this particle is larger than one picture element, i.e. greater than the distance between two adjacent scanning lines, the flying spot may scan the same particle for a second time on its next scan line. A similar sequence of events occurs, but this time the sum circuit receives the new pulse from the quantizer, plus the pulse from the previous scanning line which is released from the one-line delay unit. This state of coincident inputs produces an inhibiting pulse, which shuts the gate and prevents a second count being recorded from the same particle.

Thus in any given field every particle, however large, will be counted only once. It should be noted that this is an ideal situation and that should the profile of the particle present a re-entrant along the direction of the scan, then two counts will be recorded from the same particle. Also aggregates or particles which are not separated by more than two picture elements may be falsely counted.

The automatic sizing is performed by the sizing unit, which produces a pulse of selectable width that corresponds to a given size of particle on the sample. Any particle whose size is less than the size selected will not be recorded during a field count, so that by successive counts with selected pulses of progressively greater width (i.e. with progressively greater minimum sizes) a complete size analysis of the sample in the field can be obtained. The standard pulse widths obtained from the sizing unit are given in terms of picture element size and increase in a $\sqrt{2}$ progression; 0, 2, $2\sqrt{2}$, 4 . . . 32, $32\sqrt{2}$, 64 picture elements. The operation of the sizing unit is illustrated diagrammatically in Fig. 3. Then, considering a

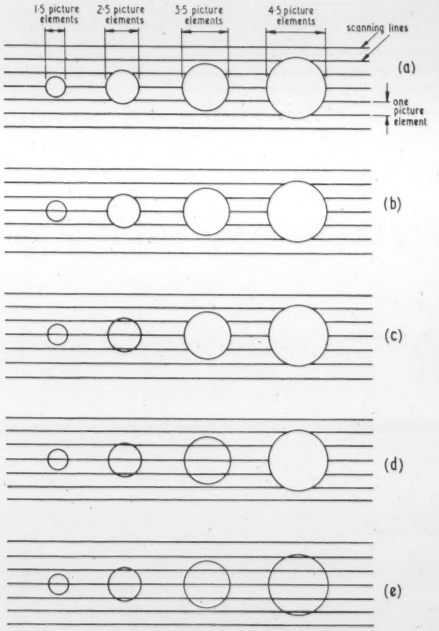


Fig. 3. Operation of the sizing unit.

(*a*) Sizing unit set on 0; no pulse is produced and therefore all four particles are counted.
(*b*) Sizing unit set on 2; the pulse produced is equivalent to two picture elements and no particle is counted whose size is less than this value. Therefore three particles are counted.
(*c*) Sizing unit set on $2\sqrt{2}$; no particle is counted whose size is less than 2·83 picture elements and therefore two particles are counted.
(*d*) Sizing unit set on 4; only one particle is counted.
(*e*) Sizing unit set on $4\sqrt{2}$; no particles are counted.

magnification such that one picture element is equivalent to 10 μ, the first count, with the sizing unit on 0, will give the total number of particles in the scanned field whose size is greater than 10 μ. The second count, with the sizing unit on 2, will give the total number of particles whose size is greater than 20 μ and so on up to the twelfth count, on 64, which will only show any particles greater than 640 μ.

On counting, one scan across the field takes 8 seconds and a further complete scan has to take place to reset the sizing unit. The total time taken to count each size range is, therefore, 16 seconds, and a complete size analysis of the field may take up to 3 minutes. Normally several fields must be counted in order to obtain a statistically sound size analysis of the whole sample; and the time taken will depend on the particle density of the sample and its degree of heterogeneity.

3. Operational characteristics

The accuracy of the instrument in counting and sizing any particular sample will depend on three main possible sources of error:

(i) errors that may be inherent in the design of the instrument itself; these include the reliability of counting and the accuracy of discrimination between the various size groups;

(ii) errors that are due to faulty operation: whilst most of the operating controls are straightforward and leave no room for error, there are some which may affect counting accuracy; these are the focusing of the microscope and the setting of the gain of the video amplifier;

(iii) errors that are due to the nature of the sample: besides the presence of aggregates or irregularly shaped particles, the accuracy may also be affected by the degree of contrast between the particles and their background and by the definition of the edges of the particles.

3.1 *Errors in design*

(*a*) *Count reliability*. In order to check the reliability of the instrument, a spray droplet sample was obtained by spraying fixed and glazed photographic bromide paper with a 1% solution of nigrosin in water. The instrument was set up on reflected light using a 25 mm objective lens and no eyepiece, giving a magnification of 50 diameters and a picture element size of 10 μ. One field of this sample was scanned repeatedly to give ten complete counts, which are shown in Table 1.

The coefficient of variation V is less than 4% in all size ranges for this particular sample and is less than 2% for all but two ranges. This indicates a remarkable degree of consistency.

One test of the instrument's reliability lies in a comparison with the reliability of a visual count. Table 2 shows ten successive visual counts made on one field of a further sample, prepared in a similar manner, and using a microscope fitted with a standard logarithmic eyepiece graticule (May 1945).

Here V increases steadily as the number decreases in the higher size groups, and it is significantly higher throughout the whole range than the variation shown by the automatic count.

Table 1. *Reliability of automatic numerical count*

Size (in picture elements)	0(1) (= 10 μ)	2 (20 μ)	2√2 (28 μ)	4 (40 μ)	4√2 (56 μ)	8 (80 μ)	8√2 (112 μ)	16 (160 μ)	16√2 (224 μ)	32 (320 μ)	32√2 (448 μ)	64 (640 μ)
Count No. 1	214	187	170	135	92	50	34	22	12	5	1	—
Count No. 2	217	192	170	140	95	48	34	23	12	5	1	—
Count No. 3	219	188	173	138	96	47	34	23	13	5	1	—
Count No. 4	215	192	174	135	96	48	35	23	12	5	1	—
Count No. 5	214	187	168	135	92	50	35	23	12	5	1	—
Count No. 6	213	191	173	140	96	51	34	23	12	5	1	—
Count No. 7	218	191	174	140	96	51	35	23	12	5	1	—
Count No. 8	215	192	174	140	96	51	34	23	11	5	1	—
Count No. 9	215	193	174	139	95	51	33	23	12	5	1	—
Count No. 10	211	188	173	139	96	50	34	23	12	5	1	—
Arithmetic mean (μ)	215·2	190·1	172·3	138·1	95·0	49·7	34·2	22·9	12·0	5·0	1·0	—
Standard deviation (σ)	2·201	2·331	2·163	2·234	1·633	1·494	0·633	0·316	0·471	0	0	—
Coefficient of variation V (%) $V = \pm 100\sigma/\mu$	1·02	1·23	1·26	1·62	1·72	3·01	1·85	1·38	3·93	0	0	—

Table 2. *Reliability of visual numerical count*

Diameter (×20 μ)	1 (20 μ)	2 (40 μ)	2√2 (56 μ)	4 (80 μ)	4√2 (112 μ)	8 (160 μ)	8√2 (224 μ)	16 (320 μ)	16√2 (448 μ)	32 (640 μ)
Count No. 1	148	90	70	41	20	11	8	4	1	—
Count No. 2	149	101	67	38	19	11	8	4	1	—
Count No. 3	151	99	66	39	19	10	7	4	1	—
Count No. 4	153	93	73	42	22	12	9	5	1	—
Count No. 5	152	100	71	41	21	12	9	5	1	—
Count No. 6	150	98	70	39	20	11	8	5	1	—
Count No. 7	155	100	74	42	21	11	9	5	1	—
Count No. 8	154	101	73	42	21	11	9	5	1	—
Count No. 9	153	98	72	40	20	11	8	5	1	—
Count No. 10	150	97	69	41	21	11	8	5	1	—
Arithmetic mean (μ)	151·5	97·7	70·5	40·5	20·4	11·1	8·3	4·7	1·0	—
Standard deviation (σ)	2·273	3·592	2·635	1·434	0·966	0·568	0·675	0·483	0	—
Coefficient of variation V (%)	1·50	3·68	3·74	3·54	4·74	5·11	8·13	10·28	0	—

In general, V would be expected to be greater in the ranges where there are relatively few particles and where an error of one count makes a larger percentage error. This is, of course, unfortunate where figures for representative particle size are to be calculated on some form of volume or mass parameter, for example, Sauter mean diameter or mass median diameter, in which case it is the groups with large diameters and (usually) small numbers which are important. Consideration of the variation in volumes of the particles in each size group in Table 2 would show larger discrepancies than those shown by the results in Table 1. Such difficulties are unavoidable and in any case are likely to be increased by sampling errors.

(*b*) *Size discrimination.* The accuracy of the instrument in sizing will be governed by the pulse lengths, which are produced by the sizing unit, and by the ability of the instrument to discriminate between particles with slightly differing diameters. To check these points, droplets were collected as before on photographic paper and the stains were sorted using a visual microscope; to reduce visual errors, the selected stains were observed by three operators. As far as possible stains were selected to give a complete progression of diameters covering the whole size range that could be counted automatically, using a 2 in. objective without eyepiece. This gave a calculated picture element size of $22{\cdot}2\ \mu$. Owing to the practical difficulty of finding suitable stains, a perfect arithmetical progression of increasing stain diameters was not possible and the actual discrimination limits, in terms of the nearest size of stain that was available, are compared in Table 3 with the theoretical limits of each size range. Column 3 shows the maximum stain size available that was not counted in the given size range and column 4 shows the next largest size available that was counted. Thus the discrimination of the instrument may be somewhat better than the results in Table 3 suggest.

Table 3. *Comparison of actual and theoretical size ranges*

Size (in picture elements)	*Calculated minimum size of particle counted* (μ)	*Actual maximum size of particle not counted* (μ)	*Actual minimum size of particle counted* (μ)
0	~22·3		20
2	44·5	50	55
$2\sqrt{2}$	63·0	71	75
4	89·0	90	95
$4\sqrt{2}$	126·0	121	129
8	178·0	172	180
$8\sqrt{2}$	252·0	245	255
16	356·0	348	357
$16\sqrt{2}$	503·0	490	510
32	711·0	705	720
$32\sqrt{2}$	1005·0	950	1020
64	1421·0	1390	1450

Except in the 2, $2\sqrt{2}$ and 4 size ranges the agreement between theoretical and actual values is very good. For a particle to be sized accurately, the clipping level should be set at the 50% level of each pulse produced when the scanning spot crosses the particle. Where the contrast is low, i.e. for small particles, this may not be achieved and consequently the particle is recorded as being slightly smaller than the actual size.

Where volume means are being used these errors are unimportant, since the volumes of particles in these ranges will usually be swamped by the volumes in the larger size groups. The magnitude of this error appears to be constant for a given type of sample so that allowance can be made. This should be done where the actual size distribution is important and where direct comparisons are to be made between different samples.

(*c*) *Count stability.* The initial warm-up period to allow the one-line delay unit to operate satisfactorily takes between 5 and 20 minutes, depending on how long the instrument has been standing idle. From the operational point of view it is obvious when this period is completed. If a count is now made and repeated at intervals of a few minutes, a slight but definite drift in the results from successive counts is found. This drift goes on for a further 10 to 20 minutes, after which the count settles down and remains constant for several hours. It is therefore desirable to switch on the instrument at least 30 minutes before the controls are finally adjusted.

After setting the controls on one sample, counts may be made on other similar samples without further adjustment. A different type of sample will alter the effective sensitivity of the instrument, and some alteration of the controls may be necessary.

(*d*) *Count accuracy.* The reliability and size discrimination of the instrument have been shown to be satisfactory when checked separately. To get some idea of how accurately it would operate during an actual size determination, a spray droplet sample was selected and an area marked out exactly equal to the area of one complete scan. Since there were only about 150 particles in this limited sample it seemed reasonable to assume that a visual count could be made with a degree of reliability similar to that shown in Table 2. Independent counts were made by three observers and the spread of results obtained is shown as the light-shaded area in Fig. 4, where they are plotted as cumulative number

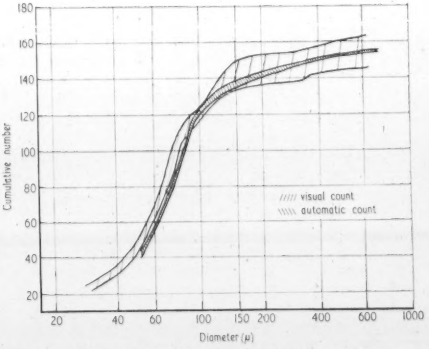


Fig. 4. Comparison of visual and automatic counts on a small sample.

against log diameter. The spread of the automatic count on this sample using 99% confidence limits is indicated within the dark-shaded area. The general agreement in size distribution between the two methods of assessment is good.

3.2 *Errors due to faulty operation of the instrument*

(*a*) *Effect of microscope focus on counting accuracy.* In order to count a statistically significant number of particles it is usually necessary to scan several fields on each sample. Initially, a particular field on the sample is focused with the high-scanning rate associated with visual operation. On

switching over to count operation this field will be counted and sized with accuracy. Movement of the microscope stage to scan another field may well upset the focus, and since the scanning rate on count operation is slow (one scan in 8 s), accurate refocusing from the count picture is difficult. This may be overcome by switching back to visual operation, checking the focus and switching over to count again. If many fields are being counted, this procedure constitutes a considerable waste of time. Within certain limits, however, it is possible to refocus on the count picture, in which case it is important to know how the counting accuracy varies when the sample is slightly out of focus.

A droplet sample was focused accurately using a 25 mm objective with a numerical aperture of 0·15 and the position of the microscope stage was recorded. By moving the microscope stage up or down, counts were made on this field at known distances above or below the plane of focus. The results are plotted in Fig. 5 (curve A) as the sum of all the

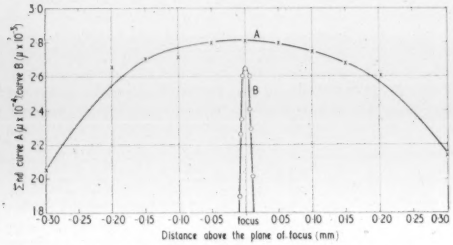


Fig. 5. The effect of mis-adjustment of the microscope focus on automatic measurement.

measured particle diameters (Σnd) against the distance above or below the plane of focus. There is little significant change in Σnd within 0·05 mm either side of the true focus, though beyond this distance the counting accuracy begins to fall away rather seriously. It is most unlikely that any normal movement of the microscope stage would upset the focus by as much as 0·05 mm. In any case, it is quite possible to refocus on the count picture to within this degree of accuracy.

The position is somewhat different when using higher-powered objectives working on transmitted light. This is shown in Fig. 5 (curve B) which is a plot of results obtained using a 4 mm objective with a numerical aperture of 0·65. This was used with a 10 × eyepiece to give a magnification of 1000 diameters and a picture element size of 0·5 μ. Here the allowable depth of focus is only about $\pm$0·003 mm and considerable practice is required to refocus to this degree of accuracy on the count picture.

(*b*) *Effect of the video amplifier gain.* The video amplifier governs both the contrast of the monitor tube picture and the amplitude of the pulses being fed to the quantizer and counting units. When the instrument is operated on count, a considerable degree of amplification is required to produce a picture on the monitor screen. As the gain is increased, a point is reached at which a picture begins to appear and as the gain is still further increased, the background of the picture steadily clears until the optimum amplification is reached. Increase in gain beyond this point produces a blurring in the picture background, an effect which can be used to visually set the video gain to a known level and thus compensate for variations in background illumination. After setting the video gain, the clipping level may require adjustment until all the particles are tagged to indicate that they are counted. The practical adjustment of the video gain requires a certain amount of practice, since very little movement of the control knob produces considerable changes in the picture on the monitor screen.

The importance of the correct setting of this control was checked by carrying out a series of counts with different settings of the gain control. This control was calibrated by a series of marks, each of which corresponded to a 3° rotation of the control knob. The total movement of the control was about 240°, but with this particular sample the picture did not begin to appear until the control had been turned through 220°, so that it is only with the last 20° that this study is concerned.

The results were obtained by using a 4 mm objective and a 10 × eyepiece giving a picture element size of 0·5 μ. One field of a sample of insecticide wettable powder was fixed under the microscope and several counts were made at various settings of the gain control. The results are plotted in Fig. 6 in terms of the sum of all the particle diameters against

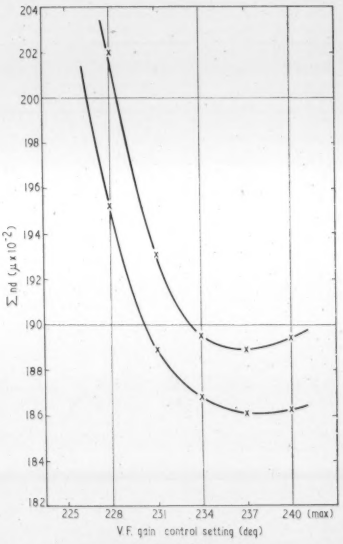


Fig. 6. The effect of video frequency gain on automatic measurement.

the rotation of the video amplifier gain control. Four or five counts were made at each setting and the curves link the maximum and minimum values obtained. The apparent size of the particles becomes smaller as the optimum gain setting is approached (237°), and above this value there is a tendency for the sum of the particle diameters to increase again. The reliability of the count is also best at the optimum setting of 237°, where the count variation is of the same order as that shown in Table 1. On both increasing and decreasing the gain, the spread of results is greater.

When allowance is made for this normal variation of results at the optimum setting of the control, it is obvious that an increase or decrease of 2° or 3° in gain control does not make

much difference from a practical point of view. If the gain is not set within these limits, considerable inaccuracies may be introduced. The ease of setting would be improved by having a fine-gain control, in addition to the coarse control.

3.3 *Errors due to sample quality*

The chief requirements in a sample for automatic counting are a reasonable degree of contrast between particle and background (this should be at least 10%), a uniform contrast across each particle (i.e. there should be no patchiness or fading towards the edges), and a clear-cut outline to the particles themselves. In the case of spray-droplet sampling the first of these requirements is easily fulfilled by adding a dye to the spray and collecting on any white or light coloured surface, whilst the others necessitate the use of a material whose surface structure is perfectly uniform and which inhibits the spread of droplets. After trying many collecting surfaces, fixed glazed photographic paper was selected as producing the best stains for use with this instrument (Furmidge, to be published); all the droplet samples that have so far been described were obtained by using this surface. The next best collecting surface was found to be high-quality glossy art paper, as used by Yeo and Coutts (1957). With this surface the definition of the edges of the stains is not so good as with photographic paper, although in all other respects it behaves in a similar manner. The effect of this poorer edge definition on count reliability is shown in Table 4, which gives the results obtained by making ten successive counts of the same field on a sample prepared by spraying 1% nigrosin solution on to glossy art paper. These results may be directly compared with the values in Table 1.

Table 4. *Reliability of counts on glazed paper*

Size (in picture elements)	0	2	$2\sqrt{2}$	4	$4\sqrt{2}$	8	$8\sqrt{2}$	16	$16\sqrt{2}$	32	$32\sqrt{2}$
Count No. 1	294	205	164	100	57	29	11	5	2	1	0
Count No. 2	296	196	157	94	53	27	13	5	2	1	0
Count No. 3	298	189	163	100	61	31	12	4	2	1	0
Count No. 4	294	197	149	103	52	32	11	4	2	1	0
Count No. 5	296	190	157	102	54	30	13	5	2	1	0
Count No. 6	279	188	154	101	56	28	13	4	2	1	0
Count No. 7	281	187	150	91	53	31	13	4	2	1	0
Count No. 8	285	192	161	103	56	32	10	4	2	1	0
Count No. 9	301	192	154	92	60	32	13	4	2	1	0
Count No. 10	290	187	158	101	55	27	12	5	2	1	0
Arithmetic mean (μ)	291·4	192·3	155·7	98·7	55·7	29·9	12·1	4·4	2·0	1·0	—
Standard deviation (σ)	7·427	5·658	5·736	4·572	2·983	2·025	1·101	0·516	0	0	—
Coefficient of variation V (%)	2·55	2·94	3·68	4·63	5·36	6·77	9·10	11·74	—	—	—

The coefficients of variation are very much larger than those calculated for the photographic paper counts and are, in fact, similar to those of the visual count shown in Table 2. As the definition of the particles on the sample deteriorates further, the reliability becomes even less. However, the results indicate that the sample quality may be somewhat removed from ideality before the realibility of automatic counting is worse than visual counting.

Besides errors due to poor definition of the particles, further errors may be introduced by overcrowding of the particles on the sample. The effect of this is illustrated in the results in Table 5. These show the percentage number frequency of stains in each size group in a spray collected on photographic paper. Sample A was an ideal sample with little or no overlapping, sample B had a denser deposit with 1–2% of the droplet stains overlapping, whilst sample C had 3–4% of the stains overlapping to some extent. After collection each sample was sized visually and automatically, the fields measured were chosen at random over the sample and just over 2000 particles were counted in each case. The counting was carried out in a rapid routine manner so that its accuracy should be similar to that obtainable in practice during a routine assessment of a number of samples.

The spray droplet spectrum was the same for each sample so that, reading across the table, the figures in each column should be the same. The visual counts show little similarity between the three samples, and a statistical analysis of the results (after allowing for the expected variation in visual counts shown in Table 2) still shows significant differences between the counts in almost every size range. Whilst it is possible that some of these errors may be due to faulty subsampling, this cannot account for the total variation between samples; in a count of 2000 particles on each sample, sampling errors should be relatively small. The time taken for each visual count was between 3 and 4 hours as against the 7 hours for similar counts given by Courshee (1954), whose results appear to be somewhat more reliable. It seems obvious that visual counts carried out over a large sample are virtually useless as a comparison standard unless they are done very slowly and carefully. Even then, as Courshee points out, their value is still dubious.

The automatic counts in Table 5 show more similarity, particularly those on samples A and B in which the agreement is very good. However, the overlapping of particles in

Table 5. *Spray droplet stains on photographic paper. (Approximately* 2000 *stains sized on each sample; time for visual count* 3–4 *h, for automatic count* 30–35 *min)*

	Number frequency (%) Visual count			*Number frequency (%) Automatic count*		
Size range	*Sample A*	*Sample B*	*Sample C*	*Sample A*	*Sample B*	*Sample C*
0	37·7	30·6	31·7	33·3	32·6	28·9
2	22·1	17·9	19·5	11·3	11·8	10·4
$2\sqrt{2}$	12·9	14·0	13·3	19·3	17·7	22·2
4	11·8	15·7	16·4	11·7	11·4	11·6
$4\sqrt{2}$	10·3	13·1	12·7	11·4	11·4	11·5
8	4·0	6·9	5·3	9·6	11·0	10·9
$8\sqrt{2}$	1·0	1·5	1·0	2·8	2·9	3·1
16	0	0·3	0·1	0·5	1·1	1·2
$16\sqrt{2}$		0·1	0	0·1	0·1	0·1
32		0		0	0	0

sample B does produce a very slight bias in the results; the proportion of particles counted in the $2\sqrt{2}$ range was significantly less than was counted in sample A, whilst the numbers counted in the 8 and 16 ranges were significantly greater. Sample C, with rather more overlapping, shows a greater bias, since the two smallest size ranges are low and there is a tendency for several of the larger ranges to give a high count. Thus, as might be expected, overlapping of particles reduces the apparent number of small ones and increases the proportion of larger ones, though the effect produced by any given aggregate depends on the profile that the aggregate presents to the line of the scanning spot.

The bias produced by aggregates, though slight, can have a serious effect on the calculation of volume or mass parameters, as is shown in Table 6 where the frequency figures for the three automatic counts have been calculated on a volume basis.

Thus an error in count of only 1 or 2% in the high size ranges can produce errors of the order of 20–30% in a calculated mass median diameter.

Table 6. *Variation in cumulative volumes due to overcrowded samples*

Size range	0	2	$2\sqrt{2}$	4	$4\sqrt{2}$	8	$8\sqrt{2}$	16	$16\sqrt{2}$
Cumulative volume A	33	123	559	1308	3373	8289	12 341	14 389	15 543
Cumulative volume B	33	127	528	1257	3322	8955	13 150	17 655	18 809
Cumulative volume C	29	112	614	1356	3437	9018	13 507	18 422	20 739

4. Conclusions

The speed and reliability of automatic measurement of microscopic particles by flying-spot scanning are considerably greater than can be achieved by visual microscopic measurement. With reasonable samples it seems probable that the degree of accuracy achieved is also considerably better, though routine visual counting is so unreliable that it cannot satisfactorily be used as a comparison standard. The speed of automatic counting enables sampling errors to be reduced, because many more particles may be measured in a given time than is possible by visual assessment. In fact, the advantages of automatic over visual counting increase considerably as the size of the samples and the number of samples to be measured increase. Even on a small sample, the automatic count is more reliable; for very large numbers of particles the accuracy of visual counting falls away rapidly, whereas with automatic counting the reliability is maintained.

The operation of the instrument is reasonably straightforward and, with practice, optimum accuracy may be readily achieved. The instrument is very versatile in that many types of samples can be measured over wide ranges of magnification, the only necessity being a reasonable degree of contrast between particles and their background. The range of magnification which may be employed with transmitted light is about 50 to 7000 diameters; with reflected light it is about 20 to 50 diameters, and for very low-power magnification the microscope may be dispensed with and samples scanned direct using a projection lens.

The measurement of particulate area on a sample using the flying-spot scanning technique is also possible; the reliability and accuracy can be expected to be similar to that achieved in size measurement.

The reliability of the instrument depends on the quality of the sample; for greatest reliability the particles should have well-defined clear-cut edges. Irregularly shaped particles which may be counted more than once will introduce some error, and lack of good edge definition will cause a considerable loss of reliability. Even so, the instrument count may still be as reliable as a visual count.

The overall accuracy with which a sample is assessed may depend to a large extent on the sampling technique used. Where the density of particles on a sample is so great that aggregation or overlapping occurs, a simple automatic scanning method is unable to discriminate between the individual particles that make up the aggregates. The magnitude of the error which is introduced depends upon the alignment of the aggregates to the line of scan, but any such error is important where volume parameters are being calculated. The ideal requirement in a sample for automatic counting is to have the maximum number of particles present per unit area without significant overlapping, the maximum number being required to obtain the greatest advantage from the speed of automatic counting.

The errors introduced by the presence of irregularly shaped particles or overlapping particles may be minimized by visually monitoring the count. The monitor screen shows all the particles as they are scanned and indicates the count pulses as they are recorded. This fact enables a visual assessment to be made of the instrument's operation and in this way it is possible to maintain a high degree of accuracy on all but the most overcrowded samples.

Acknowledgments

The author would like to thank Prof. H. G. H. Kearns for his advice and encouragement and Mr. T. E. Cobbald for many useful discussions. Thanks are also due to several of the staff of Messrs. Rank Cintel Ltd., in particular to Mr. D. Causley for advice and assistance in the operation of the 'Flying-spot particle resolver', and to Miss D. I. Conibear and Miss G. Browning who carried out much of the practical work.

Figures 1 and 2 are reproduced by permission of Messrs. Rank Cintel Ltd.

References

CAUSLEY, D., and YOUNG, J. Z., 1955, *Research*, **8**, 430.

COURSHEE, R. J., 1954, *Brit. J. Appl. Phys., Suppl. No.* 3, **5**, *S.* 161.

HAWKSLEY, P. G. W., 1954, *Brit. J. Appl. Phys., Suppl. No.* 3, **5**, *S.* 125.

HAWKSLEY, P. G. W., BLACKETT, J. H., MEYER, E. W., and FITZSIMMONDS, A. E., 1954, *Brit. J. Appl. Phys., Suppl. No.* 3, **5**, *S.* 165.

MAY, K. R., 1945, *J. Sci. Instrum.*, **22**, 187.

MORGAN, B. B., and MEYER, E. W., 1959, *J. Sci. Instrum.*, **36**, 492.

TAYLOR, W. K., 1954, *Brit. J. Appl. Phys., Suppl. No.* 3, **5**, *S.* 173.

YEO, D., and COUTTS, H. H., 1957, *Nature, Lond.*, **179**, 864.,

Gaseous diffusion in porous media. Part 3—Wet granular materials

by J. A. CURRIE, Ph.D., Rothamsted Experimental Station, Harpenden, Herts.

MS. received 27th August 1960

Abstract

The diffusion of hydrogen through granular materials partly saturated with water was measured by a non-steady state technique previously described. The two types of sample used consisted of solid particles (unimodal pore-size distribution) and porous particles (bimodal pore-size distribution), and all measurements were made on samples being dried from saturation. Coefficients of diffusion D were calculated, and for the range over which the larger pores were emptying the empirical equation $D = D_v(\epsilon/\epsilon_v)^\sigma$ *fitted all materials, where* ϵ *is the fractional air-filled volume,* ϵ_v *is the volume occupied by the larger-pore phase, and where* D_v *is the diffusion coefficient when only this phase is air-filled. For all materials* $\sigma \simeq 4$ *whether the samples were uniform or of mixed sizes. No such relationship existed over the subsequent range in which the smaller pores were drained. Over this range D must be a function of at least five independent variables— the total porosity* ϵ_T*, the crumb porosity* ϵ_c*, the shape factor for the crumbs or inter-crumb pores k, the shape factor for the particles forming the crumbs or crumb pores* k_c*, and the moisture content of the sample. The spatial distribution of pores within a porous medium can be as important as the sizes of the pores. The factors k and m, previously introduced as particle-shape factors, now have a greater significance as measures of the geometrical complexity of a porous system. Adding water can either increase or decrease the complexity, depending on the amount added and the nature of the system. The agricultural significance of diffusion between the crumbs* D_v *and within the crumbs* D_c *is discussed, and it is suggested that* D_c*, or its associated complexity factor* k_c*, might be used as an index of soil structure.*

Introduction

DECREASING the air-filled pore space in a porous medium decreases the rate at which gases will diffuse through it. Though adding water is a convenient way of obtaining a continuously variable pore space, there is little evidence to show whether it acts differently from other ways of decreasing pore-space, such as, for example, repacking, or mixing particles of different sizes. Nor is there much evidence to show whether diffusion through different types of porous media is influenced in the same way by wetting. Hannen (1892) used a non-steady state method to measure carbon dioxide diffusion through dry and partly saturated packings of soils and sands. His data, re-calculated to allow for the exponential rate of decay of carbon dioxide concentration in the non-steady state (from the data available this can only be approximate), show for dry materials a relationship of the form $D/D_0 \simeq 0{\cdot}6\,\epsilon$. All the points for wetted materials lie below this line when graphed, suggesting that, in his experiment, moisture was more effective than packing in decreasing diffusion. Buckingham (1904) used a steady state technique on both wet and dry sands and soils. From his data for carbon dioxide, the relationship $D/D_0 = \epsilon^2$ emerged, but his range of moisture contents was too small to establish the role of moisture, other than as a pore filler. Soils were the only materials to which Penman (1940) added moisture, wetting them before they were packed into the diffusion cylinder, rather than adding water to already packed soils. His values of D/D_0 for moist soils do not differ from those for other (dry) materials packed to the same porosity. Hagan (1941) measured diffusion rates through soils wetted with up to 18% moisture, and concluded that moisture content affected the permeability to carbon bisulphide more than any other variable studied. By extrapolating regression lines for their diffusion data, Blake and Page (1948) showed that, where diffusion through one soil ceased at a porosity of 0·1, in another it reached zero only when the porosity was zero. They suggested that this may have resulted from the blocking of some of the air-filled pore-space by water films, but they did not give moisture contents. Taylor (1949) measured diffusion through a loam soil, a quartz sand and through powdered glass, interpreting his data in terms of 'an equivalent diffusion distance' (actually this is equivalent to $(D/D_0)^{-1/2}$). This parameter bore no simple relationship to moisture content, to moisture potential or to air content, and the effect of moisture on diffusion was obscured. Measurements on soils *in situ* by Raney (1949) have too few supplementary data for the effects of cultivation to be separated from the effects of moisture. Bruce and Webber (1953) used Raney's method on a loam soil and suggested that the relationship between diffusion and moisture content was closely related to pore-size distribution; they also stressed the inadvisability of extrapolating for diffusion rates at moisture contents outside the experimental range. Call (1957) measured the diffusion of ethylene dibromide through soils packed into tubes to simulate field conditions and his results supported $D/D_0 = 0{\cdot}66\ (\epsilon - 0{\cdot}1)$. He too stressed the importance of 'blocked pores' but, with no data for moisture contents, the effect of moisture is difficult to separate from effects of dry porosity, soil condition, or possibly anisotropy of structure induced by compacting the soils. Rust *et al.* (1957) wetted non-soil particles, and found $D/D_0 = 0{\cdot}68\epsilon - 0{\cdot}01$, whereas for the dry materials $D/D_0 = 0{\cdot}60\epsilon + 0{\cdot}04$. Their results suggest that moisture-blocked pores decrease diffusion more than pores blocked in other ways, and the values for wet soils lie close to the theoretical curve for cylindrical particles with blocked pores (De Vries 1950) and in no case above the curve for spherical particles. In an analogous study of the electrical conductivity of electrolyte solutions saturating various granular

materials, Wyllie and Gregory (1953) investigated the effect of depositing a non-conducting compound between the particles. This they put into position in the dry packing in solution. The deposit of silica, which remained when the solvent was removed by evaporation, occupied the same sites as water would be expected to occupy when added to the same type of packing. If the analogy is exact, their results suggest that the relation $D/D_0 \propto \epsilon^{4\cdot2}$ might be expected.

Theoretical investigations of the type used for dry materials are few. The equation $D/D_0 = \epsilon^{3/2}$ of Marshall (1959) makes no distinction between wet and dry materials. Millington (1959) modifies his relationship for dry soils ($D/D_0 = \epsilon^{4/3}$) to read $D/D_0 = n^2(\epsilon_1^{4/3}/m^2)$, where m equal-volume pore-size groups make up the total porosity ϵ, and when n of these are drained, the air-filled pore-space is given by $\epsilon_1 = n(\epsilon/m)$. The effect is to predict a lower rate of diffusion through moist materials than through dry materials of equal air-filled porosity. De Vries (1950) has suggested a 'blocked-pore' factor for use with his equation, but no existing method distinguishes at a given moisture content between pore-space which is merely water-filled and pore-space which might be termed 'blocked'.

Qualitatively the presence of water seems more effective in decreasing gaseous diffusion than are other agents that decrease air-filled pore-space to the same extent, but quantitatively it is impossible at present to predict the decrease in diffusion rate for a given type of material.

Measurement of diffusion

The apparatus described in Part I (Currie 1960a) was used to measure the diffusion of hydrogen through the porous samples. To use the technique of Part II (Currie 1960b), it was first necessary to overcome difficulties not encountered with dry materials.

The supporting gauze, as previously, had to be rigid, able to retain the granules, and offer the minimum resistance to diffusion. In addition, with moist samples, the apertures had to remain open in diffusion experiments, but had to be able to transmit water as the sample was brought to moisture equilibrium on a suction plate. Three attempts to avoid the use of permeable supports failed.

In the first, granular packings were stabilized by adding adhesive in solution and then allowing it to dry, but the large quantity needed to produce stability raised doubts about the nature of the resultant pore system (Wyllie and Gregory (1953) showed the effect of adding a cementing phase to various granular materials). Similarly, sintered-glass blocks were thought to be atypical of unconsolidated packings.

In the second attempt, a U-shaped tube of square section was used—copying van Bavel (1952)—with one end fitted into the sample tube recess and the other open to the atmosphere. Though this supported the porous material without in any way impeding 'longitudinal' gas flow, it had two disadvantages. In the second limb, the upward diffusion of hydrogen was assisted by convection to an unacceptable extent; and with some wet materials, a moisture gradient was set up within the sample, which was capable of sealing the bottom of the U-tube while leaving the sample as a whole unsaturated.

In the third attempt, the lower end of the straight cylindrical tube was sealed and hydrogen from the gas tube diffused to a concentration in equilibrium with the air in the porous sample. Though a simple working equation was developed for this system, in practice the interval over which the changing concentration could be followed was too small to allow D to be reliably estimated.

Because these attempts failed, it was necessary to resort to the permeable support. By careful choice of suitable open-meshed nylon fabrics, supports fulfilling all requirements for specific materials could be provided.

The second new difficulty was the variability of moisture distribution throughout the samples. When samples packed in 3 in. tubes were nearly saturated, some downward drainage occurred. At best, a measured diffusion coefficient could then be related to a mean value of saturation but diffusion through the sample often ceased when only the lower end of the sample was blocked by moisture. Shorter tubes (1 in.) were satisfactory, and when made of Perspex allowed the presence of moisture gradients to be observed. Occasionally, the visual checks were supplemented by destructive testing. To maintain the convenient time scale of experimental operations, the cross-sectional area of the tubes was altered from approximately 22 cm^2 to about 8 cm^2. These tubes were fitted with flanges at the upper ends to enable them to fit snugly into the apparatus in place of the standard brass tubes.

The wet samples required careful preparation to give reproducible diffusion coefficients. All the measurements were made on samples being dried from saturation. After an initial diffusion measurement, when dry, the sample was evacuated in a desiccator; air-free distilled water was then admitted and allowed to saturate and almost submerge the sample before the whole was restored to atmospheric pressure. The total pore-space calculated from the weight of liquid held at saturation agreed well with that derived from the weight and density of the solid phase, indicating that saturation was complete. The sample was then placed on a suction plate (sintered glass or other porous material), brought into equilibrium with a chosen water potential, a set of diffusion measurements made, and the moisture content determined by weighing. A series of such operations at increasing moisture tensions was continued until the suction plate technique became inadequate. When further increases in tension failed to remove enough water, the remainder was removed slowly, either by passing a stream of air through the sample, or by bringing the sample to equilibrium over saturated solutions of selected salts. A final diffusion measurement was made after the sample had been completely dried over phosphorus pentoxide. Comparison of initial and final diffusion coefficients for the dry sample showed whether the particles had been obviously re-orientated in the sample during handling. In practice, agreement between the coefficients was always well within the 1 % tolerance permitted between duplicate determinations on the same sample.

In samples prepared by desaturation, the continuity of air-filled pore-space with the outside atmosphere was ensured. Saturating dry samples by small increments of water often entrapped air, isolating pockets that could not contribute towards gaseous diffusion through the sample. For this reason, which must not be regarded as a hysteresis effect, reproducible results were obtainable only on the drying cycle.

Materials used. The materials used were in two main groups—uniform solid particles and uniform porous particles. The first group comprised glass spheres, sand and carborundum and, for convenience, the sintered-glass blocks (Porosity 2); the second included pumice and soil crumbs. Three soils were used, chosen as having different crumb porosities, ϵ_c (i.e. the fractional volume of the crumb free from solid matter), and expected to have differing internal structure. Two of these soils were unstable in water and all three swelled on wetting. Though this swelling (and shrinking) doubtless has an important influence on the diffusion

of gases and ions, particularly in agricultural soils, it was considered desirable to exclude it from the present work. The soils were therefore heated to 800° C, which stabilized the crumbs, increased crumb porosities by from 6 to 9% of the crumb volume, and probably modified the original structure slightly.

Experimental results

With all the materials tested, adding water caused a decrease in the diffusion coefficient (Fig. 1), but only in packings of

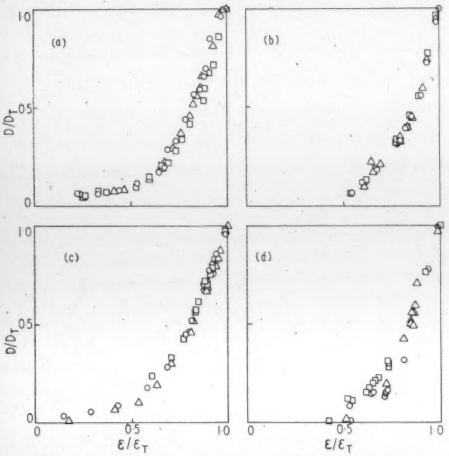


Fig. 1. Relative diffusion/porosity relationships for uniform solid particle systems on drying from saturation (left to right): (*a*) sintered-glass blocks; (*b*) glass beads; (*c*) sand; (*d*) carborundum.

Sample 1, □; Sample 2, ○; Sample 3, △.

solid particles was this consistently greater than that caused by equal decreases in porosity achieved by repacking or particle mixing. When plotted on a log–log basis (Figs 2 and 6) the data for all materials may be fitted over a substantial (lower) range of porosity by $D/D_0 = \eta\epsilon^\sigma$ where η is a function of the total porosity ϵ_T and of the corresponding diffusion ratio D_T/D_0 for solid particles, or of the inter-crumb porosity ϵ_v (Fig. 9) and corresponding diffusion ratio D_v/D_0 for porous particles. For both types of particle $\sigma \simeq 4$. When there is any deviation from this equation over the lower range, it occurs as $\epsilon \to 0$ in the sense that $\sigma < 4$, whereas theories involving 'blocked pores' and 'cut-off' values at low porosities lead to $\sigma > 4$. Particle shape influences the rate of diffusion through dry materials (Currie 1960b), but from Figs 1 and 2 it is apparent that in wet materials particle shape has much less effect on diffusion, for fairly obvious reasons. For the former, particle shape is the main determinant of pore shape, but in wet materials the pore shape is very much modified by the presence of water. Whatever the shape of the empty pore, the water menisci between the particles 'round off' the boundaries of the air-filled volume, so that at a given degree of saturation, most solid-grain packings will exhibit the same pore shape. Exceptions to this are inevitable, and for particles of irregular shape a higher degree of saturation is expected to be needed before the pore shape ceases to be characteristic of the particles. Packings of porous particles are in this category, for they must be completely saturated

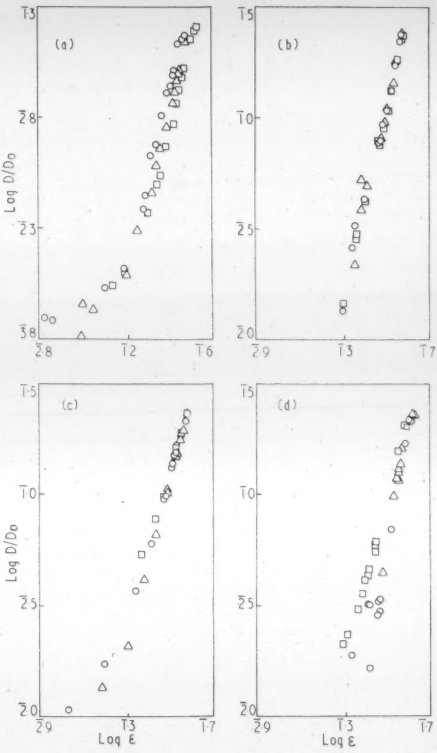


Fig. 2. Logarithmic diffusion/porosity relationships for uniform solid particle systems: (*a*) sintered-glass blocks; (*b*) glass beads; (*c*) sand; (*d*) carborundum.

Sample 1, □; Sample 2, ○; Sample 3, △.

before they become representative of the solid particle packing. Figs 5 and 6 show this and will be discussed later.

Figures 1 and 2 refer to materials of uniform particle size having sigmoid moisture retention curves (Fig. 3, curves (*a–d*)) characteristic of a normal pore-size distribution. To show that it is the normal distribution of pore size and not uniformity of particle size which governs this type of moisture/diffusion relationship, the measurements were repeated on a mixture of 5–6 and 0·75–0·8 mm diameter spheres, and on mixtures of 0·5–1·0 and 1–2 mm crumbs for one of the soils. Moisture/diffusion curves (Figs 4 and 8) and moisture retention curves (Fig. 3 and Fig. 7, curve (*e*)) for these mixtures are similar to those obtained for particles of uniform size.

Porous particles give packings with at least two distinct kinds of pores, namely those between the particles, the *inter*-crumb pores, and those within the particles, the crumb pores

(Fig. 9). (Orr and Dallavalle (1959) use the terms 'voids' and 'pores' respectively for these two groups.) Such a system has discrete zones of smaller pores separated by a continuous

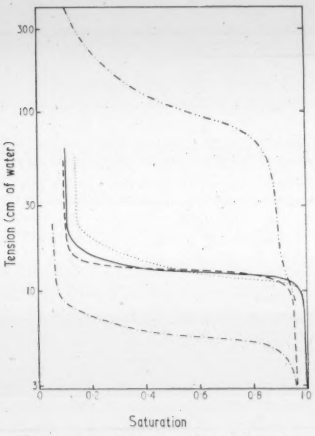


Fig. 3. Moisture retention curves for solid particle systems.
(*a*) Sintered glass
(*b*) Glass beads
(*c*) Sand
(*d*) Carborundum
(*e*) Mixed glass beads

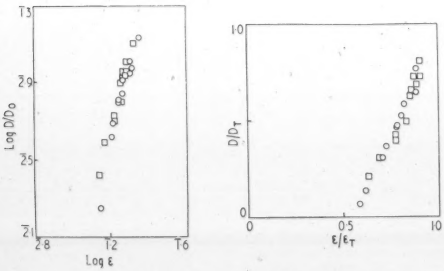


Fig. 4. Diffusion/porosity relationships for mixed solid particles.
Sample 1, □; Sample 2, ○.

system of larger pores, and is characterized by a 'double sigmoid' type of moisture retention curve (Fig. 7—only the lower sigmoid is shown). Desaturation/diffusion curves for such samples (Figs 5 and 6) reveal two distinct phases corresponding to desaturation of, first the inter-crumb pores, and second the crumb pores. In the first phase, the process is identical with the desaturation of the simple pore system, the saturated granules behaving exactly as solid particles of the same shape and volume. (Four systems were fully investigated, and limited experiments with ten similar materials produced no evidence contrary to this finding.) In the second phase, there is no consistent relationship between the volume of the crumb pores and the increase in diffusion brought about when they are drained. This will be better

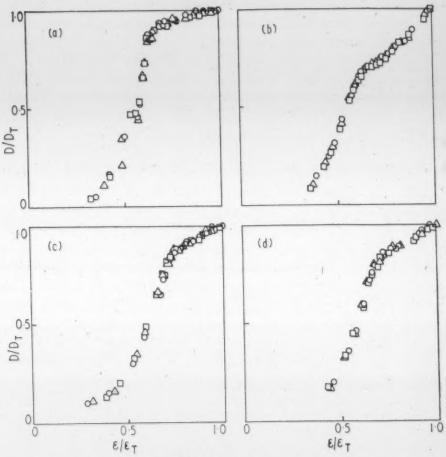


Fig. 5. Relative diffusion/porosity relationships for uniform porous particle systems. (*a*) Pumice; (*b*) HPP soil crumbs; (*c*) Ba80 soil crumbs; (*d*) TBL soil crumbs.
Sample 1, □; Sample 2, ○; Sample 3, △.

understood with reference to the crumb material. The equation of De Vries (1950) for a dry packing of porous crumbs may be written as

$$\frac{D}{D_0} = \frac{1 - (1 - kD_c/D_0)(1 - \epsilon_v)}{1 + (k - 1)(1 - \epsilon_v)}$$

where k is a shape factor for the crumb, and where ϵ_v is the inter-crumb porosity. D_c/D_0 represents the diffusion ratio for a continuous expanse of crumb material and will itself be a function of the crumb porosity ϵ_c, and the shape of the particles forming the crumb (shape factor k_c). As the crumb material is a consolidated medium, k_c will have a wider range of values than for unconsolidated media and will confer on D_c/D_0 an extended lower range of values. At the beginning of phase two, when the inter-crumb pore-space is completely empty, and the crumb pore-space is completely full $D_c = 0$ and at this stage

$$\frac{D}{D_0} = \frac{\epsilon_v}{1 + (k - 1)(1 - \epsilon_v)}$$

i.e. the saturated crumbs act as solid particles of the same shape and volume. As phase two continues, water is withdrawn from the crumbs, and D_c becomes positive as the total air-filled pore-space increases to include a crumb component. Search for a simple power law relating D_c and crumb pore-space has so far been unsuccessful. The value of D_c must depend on the porosity of the crumb, the fraction of this occupied by water, and the shape of the particles making up the crumb, so that the measured diffusion ratio D/D_0 must be dependent on at least five independent variables, for to these three have to be added the inter-crumb pore-space, and the shape factor for the crumbs as units. Not much is known about the internal geometry of the crumbs, but the crumb porosities ϵ_c are given in the Table. The ease with which the different crumbs were saturated suggests that the

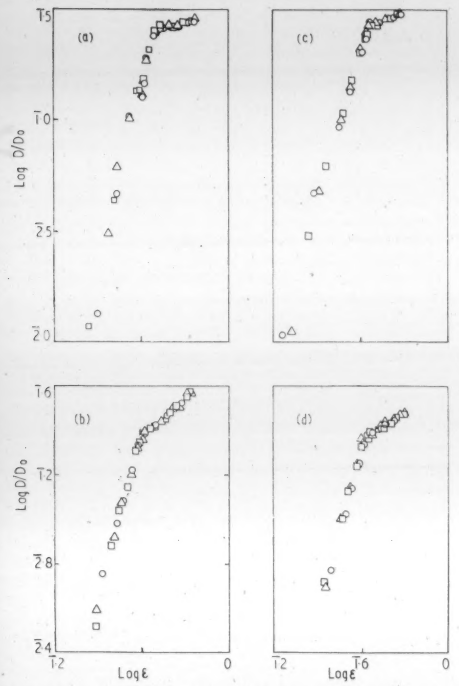


Fig. 6. Logarithmic diffusion/porosity relationships for uniform porous particle systems. (*a*) Pumice; (*b*) HPP soil crumbs; (*c*) Ba80 soil crumbs; (*d*) TBL soil crumbs.

Sample 1, □; Sample 2, ○; Sample 3, △.

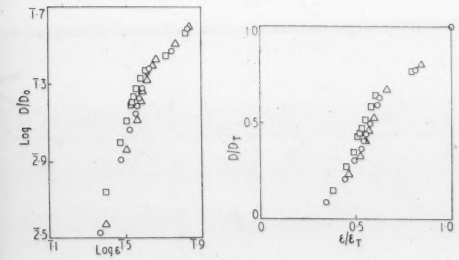


Fig. 8. Diffusion/porosity relationship for mixed porous particles.

0·5 — 1·0 : 1·0 — 2·0

Sample 1	1	:	2	□
2	2	:	1	○
3	3	:	0	△

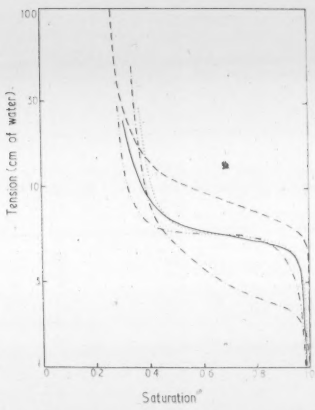


Fig. 7. Moisture retention curves for porous particle systems.

(*a*) Pumice
(*b*) HPP soil crumbs
(*c*) Ba80 soil crumbs
(*d*) TBL soil crumbs
(*e*) Mixed HPP soil crumbs

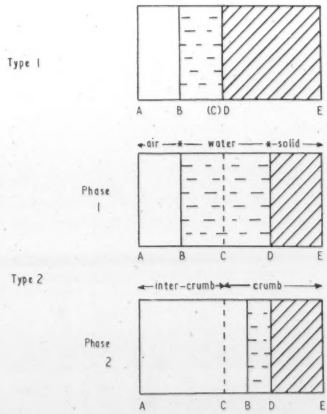


Fig. 9. Diagrammatic pore distribution for solid particle systems (type 1) and porous particle systems (type 2).

1. Expressed as a fraction of total volume

		Corresponding diffusion coefficient
Total pore-space	ϵ_T = AD/AE	D_T
Inter-crumb pore-space	ϵ_v = AC/AE	D_v
Crumb pore-space	ϵ_p = CD/AE	—
Air porosity	ϵ = AB/AE	D

2. Expressed as a fraction of crumb volume

Crumb porosity	ϵ_c = CD/CE	D_c

internal geometry of the HPP crumbs was least complicated and that of the pumice most complicated. The agricultural history of the soils suggest that the HPP crumbs have a better internal structure than either the Ba80 or TBL crumbs, the first of which are from a continuously cultivated arable soil, and the second from a boulder-clay sub-soil in which a high degree of particle orientation might be expected.

Table. *Data for dry samples*

	ϵ_T	D_T/D_0	ϵ_c
	0·303	0·137	
Sintered glass	0·301	0·148	—
	0·342	0·162	
	0·378	0·236	
Glass beads	0·373	0·237	—
	0·376	0·239	
	0·387	0·230	
Sand	0·379	0·214	—
	0·381	0·224	
	0·386	0·204	
Carborundum	0·408	0·216	—
	0·419	0·228	
	0·701	0·285	
Pumice	0·703	0·276	0·517
	0·700	0·280	
	0·673	0·378	
HPP	0·676	0·375	0·418
	0·677	0·374	
	0·612	0·307	
Ba8O	0·617	0·306	0·350
	0·612	0·306	
	0·633	0·300	
TBL	0·638	0·298	0·295
	0·634	0·301	
	0·213	0·128	
Bead mixtures	0·226	0·136	—
Soil mix 1 : 2	0·656	0·363	
2 : 1	0·674	0·381	0·418
3 : 0	0·686	0·393	

Only two of an almost infinite number of types of porous system have been considered here, but they are those most commonly encountered in practice. Pore-size distribution is of apparent importance, but other materials may have similar pore-size distributions without necessarily exhibiting the same moisture/diffusion characteristics, or the same moisture retention curves, for example one in which the larger pores are isolated in a continuous matrix of smaller pored material. It is therefore necessary to specify also the spatial distributions of these pore-size groups and the continuity of the pores. Moisture retention curves seem the necessary compromise between these three requisites.

In Part II, shape factors, k and m were calculated for all the dry materials used, recognizing that the values for the porous particles were 'effective' shape factors. Values for k and m were calculated for one sample from each of the four different materials of this type (Fig. 10). Both pass through a sharp minimum at the water content when only the crumbs are saturated, showing a very clear transition from phase 1 to phase 2 and these values represent the factors for the true shape of the crumb. The values for the four materials differ in a way consistent with the observed differences in particle shape. The factors k and m, originally introduced as particle shape factors, can now be given greater significance as measures of the complexity of the pore system, a concept equally applicable to consolidated and unconsolidated materials. Starting from dry materials and adding water (moving from right to left in Fig. 10), there is a steady decrease in complexity as the crumb pores are filled, reaching the minimum at crumb saturation. Further additions begin to fill spaces between crumbs, increasing the tortuosity of the remaining air path and so increasing the complexity again.

Discussion

It is apparent for a given porous material that, unless certain geometrical properties are known, the effective

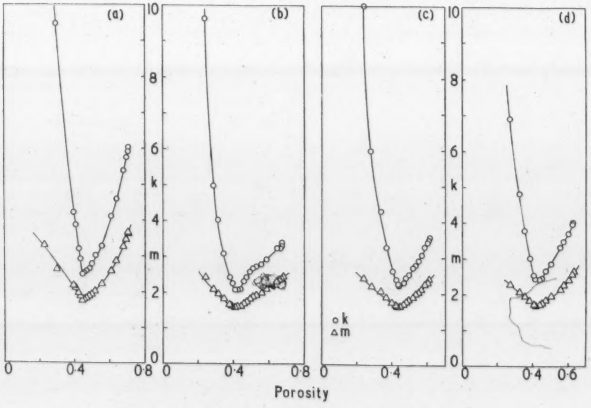


Fig. 10. The effect of moisture on 'shape factors'.
(*a*) Pumice; (*b*) HPP soil crumbs; (*c*) Ba80 soil crumbs; (*d*) TBL soil crumbs.

diffusion coefficient cannot be interpreted in terms of moisture content. Further detailed examination of published data reviewed earlier must be restricted to those materials for which these specifications appear complete. Taylor's data for quartz sand can be fitted with a fourth power relationship but not those for powdered glass. In assessing the other somewhat contradictory data, it must be remembered that adding water can simplify or complicate the system for diffusion according to type of medium. Values for k and m calculated from Penman's data are lower for wet than for dry soils, consistent with what is known of the method of packing. Call's data are not inconsistent with a fourth-power relationship between diffusion and porosity, bearing in mind that the packings were not of the same total porosity or all of the same crumb porosity, factors ignored in fitting a linear relationship. Other results interpreted by their authors in terms of 'cut-off' values, may be satisfactorily explained by $D/D_0 \propto \epsilon^4$, for it is not always possible to distinguish between first and higher power relationships, especially over limited range of porosity and with a high experimental error. The relation $D/D_0 \propto \epsilon^{4\cdot2}$, suggested from the work of Wyllie and Gregory, indicates that the analogy between cement and water is correct. Millington's equation by comparison gives $D/D_0 \propto \epsilon^{10/3}$ for wet materials.

Most agricultural soils can be placed in the two categories studied here. Soils where sand predominates are of the first type. Most soils, however, contain some clay and because their primary particles aggregate into discrete units separated by larger pores, they are of the second type. These units are formed either by natural processes (wetting and drying, freezing and thawing) or by artificial processes (cultivation); they are the blocks, prisms, plates, etc., of the undisturbed sub-soil, and the clods, crumbs and aggregates of the cultivated layer. The larger inter-crumb pores permit surplus water to be removed easily and allow gases to diffuse between the 'crumb' surface and the soil surface; the smaller crumb pores hold the reserve of water for plant growth and are the site of much of the biological activity in the soil. For this activity to proceed satisfactorily, i.e. in a direction beneficial to crop growth, oxygen must reach these smaller pores and carbon dioxide must be removed, both gases diffusing through partly saturated crumb pores. The interchange of gases between respiring organisms and roots in the soil involves two of the factors studied here, firstly D_v/D_0 for the diffusion in the inter-crumb pores, and secondly D_c/D_0 for the diffusion between the crumb surface and the respiring surface. In addition to gaseous diffusion within the crumbs, there is also ionic diffusion induced by nutrient uptake by the organisms. By analogy, D_c/D_0 will be the same for ions in the saturated crumb as it is for gases in the dry crumb, but the same power law will not necessarily govern both at intermediate saturations. The value of D_c/D_0 or its associated complexity k_c is likely therefore to be an important parameter for soil structure, applicable both to individual crumbs and to larger units in an isotropic soil. Though the full relationship has yet to be determined, the necessary parameter can probably be calculated from the values of ϵ_c, D_v and D_T obtained from the methods described.

Acknowledgments

The author again wishes to thank Dr. H. L. Penman for encouragement and valued discussion. He also wishes to thank Mr. M. R. Richardson for making the numerous diffusion measurements.

References

VAN BAVEL, C. H. M., 1952, *Soil Sci.*, **73,** 91.
BLAKE, G. R., and PAGE, J. B., 1948, *Soil Sci. Amer. Proc.*, **13,** 37.
BRUCE, R. B., and WEBBER, L. R., 1953, *Canad. J. Agric. Sci.*, **33,** 430.
BUCKINGHAM, E., 1904, *U.S. Dept. Agric. Bureau of Soils Bull.* No. 25.
CALL, F., 1957, *J. Sci. Food Agric.*, **8,** 143.
CURRIE, J. A., 1960a, *Brit. J. Appl. Phys.*, **11,** 314.
—— 1960b, *Brit. J. Appl. Phys.*, **11,** 318.
DE VRIES, D. A., 1950, *Trans Fourth Int. Congr. Soil Sci.*, **2,** 41.
HAGAN, R. M., 1941, *Hilgardia*, **14,** 83.
HANNEN, F., 1892, *Forsch. Gebiete Agr. Phy.*, **15,** 6.
MARSHALL, T. J., 1959, *J. Soil Sci.*, **9,** 1.
MILLINGTON, R. J., 1959, *Science*, **130,** 100.
ORR, C., and DALLAVALLE, J. M., 1959, *Fine Particle Measurement* (New York: Macmillan).
PENMAN, H. L., 1940, *J. Agric. Sci.*, **30,** 570.
RANEY, W. A., 1949, *Soil Sci. Soc. Amer. Proc.*, **14,** 61.
RUST, R. N., KLUTE, A., and GEISEKING, J. E., 1957, *Soil Sci.*, **84,** 453.
TAYLOR, S. A., 1949, *Soil Sci. Soc. Amer. Proc.*, **14,** 55.
WYLLIE, M. R. J., and GREGORY, A. R., 1953, *Trans Amer. Inst. Min. Eng.*, **198,** 103.

Electrode contamination and arc formation

by H. de B. KNIGHT, M.Sc., M.I.E.E., F.Inst.P., A.E.I. (Rugby) Research Laboratory, Rugby

MS. received 21*st December* 1960, *in revised form* 1*st March* 1961

Abstract

When the surface of the cathode of a gas discharge is contaminated with insulating substances or with mercury, the discharge may develop into an arc with a lower voltage drop than when the contaminant is not present. Insulation particles are added to the electrodes of spark gaps designed to have a low impedance. With mercury, the minimum voltage and current values required for arc formation are lower the more effectively the mercury wets the electrode surface. Both contaminants play a part in the phenomenon of arc-back in arc rectifiers. The probability of arc-back depends on the random concentration of ions and also on the nature and extent of the contamination of the anode; this varies both with the conditions of operation and with previous treatment. The paper discusses the interpretation of tests made in 'synthetic' circuits to determine the probability of arc-back in a rectifier in service.

1. Theories of arc-back

THE problem of arc formation in a plasma is typified by that of arc-back in rectifiers. It is commonly accepted that insulating particles on the anode surface are a major cause of arc-back, as originally suggested by Langmuir (1928). At the end of the forward conduction period the insulating particles become charged by ions from the residual plasma, these being accelerated towards the anode by the field due to the inverse voltage; if the potential of any particle rises sufficiently above that of the adjacent conducting surface, the local electric field causes field emission of electrons, and this may lead to the formation of an arc.

The theory is consistent with numerous experimental observations. Malter (1936) observed electron emission from an oxidized aluminium cathode when the oxide layer became charged with positive ions. Kingdon and Lawton (1939) made tests in which an electrode immersed in a plasma was made momentarily negative by the discharge of a capacitor; they found that the voltage required for arc formation was considerably lowered if the electrode surface was contaminated by insulating particles, such as powdered quartz, alumina or glass, and that the effect disappeared when the insulation resistance of the particles was reduced by heating the electrode to about 1000° C.

Earlier experimenters (von Issendorff, Schenkel and Seeliger 1930, Dällenbach, Gerecke and Stoll 1925, von Issendorff 1929, Slepian and Ludwig 1932), studying the problem by applying potentials to electrodes immersed in plasma, as well as by a wide variety of inverse-current and arc-back tests in rectifiers, had noted that an arc is liable to be formed when positive ions, accelerated towards the anode, fall on foreign substances on its surface. It was thought that the initiating mechanism was thermal emission from particles heated by ion impact. This could be the case if the contamination were of low work-function material, especially if it were in such loose contact with the surface as to be easily raised to a high temperature.

It is likely that both thermal and field emission agencies contribute to arc-back; another agency which does not obviously fit either theory, and to which further attention is given below, is the presence of mercury on the anode.

2. Arc-back probability

Kingdon and Lawton (1939) studied the validity of the insulating-patch theory by observations of arc-back in mercury-arc rectifiers. On this theory, the probability of arc-back would depend on the chance of an adequate concentration of ions at a particular spot at a critical moment. If the distribution of ions in the space close to the anode is assumed to be random, the chance may be expressed by the Poisson probability formula: if the average number of ions collecting at a spot of given size in a given time is m, then the probability that in any individual trial the number will be n is expressed by

$$P(n) = \frac{m^n \epsilon^{-m}}{n!}.$$

If n is the number leading to arc-back, $P(n)$ denotes the arc-back probability. A family of curves may be drawn showing $P(n)$ as a function of m for different values of n.

In the present case neither m nor n is known. The average value of m depends on the number of ions available and on their rate of arrival at the sensitive spots. Kingdon and Lawton assumed it to be proportional to di/dt, the final rate of fall of the current (which represents the number of residual ions in the space at the end of the conduction period) and to V, the initial inverse voltage; the factor V was adopted 'for a trial' as representing the rate of flow of the ions towards the anode. Their experimental observations of arc-back probability were then plotted against an arc-back factor $F = Vdi/dt$, the points being fitted to the probability curve to which they made the best match. This gave a value for n (which was a characteristic of the curve selected) and also for the constant k in the assumed relation $m = kF$. The number n required for arc-back varies with the state of ageing of the device; it was assumed that the value of k obtained with a fully aged rectifier, that is when arc-back is due to the most sensitive spot left permanently on the anode, applies as a constant for the device (although dependent on pressure).

The work described was intended mainly to test 'a particular hypothesis' relating arc-back to the charging of insulating particles. The factor F has been generally accepted as a useful working guide for circuit design, and particularly

as indicating a maximum condition not to be exceeded. Later writers (Steiner, Zehner and Zuvers 1944) proposed to extrapolate the probability curves so as to predict the behaviour of a rectifier in service; this is discussed later.

3. The development of the arc

The conditions at the anode surface may give rise to a reverse current without producing arc-back. The transition to a self-sustaining arc regime requires an agency which facilitates electron emission, not only initially but also during the development of the discharge. With insulating particles this condition is fulfilled because the particles retain their charge as new ions formed in the discharge replace those which leak away.

Further, the current must reach the minimum value required for cathode-spot stability while the favourable conditions still obtain. The inductance of the circuit sets a limit to the rate of rise of current; it also tends, however, to reduce the value of current required for arc-spot stability. In an experimental test with a mercury-pool rectifier, the cathode spot for a current of 2·5 A had a most probable life of about one minute with negligible inductance in circuit; but the arc was stable throughout an hour's test at 1·2 A when an inductance of 7 mH was connected in series.

It is common practice to insert reactance in rectifier circuits in order to limit the possible fault current to a maximum permissible value. This limits both the rate of fall of forward current and the rate of rise of inverse current, the two rates being identical if arcing occurs. A reduction in the rate of current change reduces the danger of arc-back in two ways: it reduces the residual ionization and also increases the chance that the conditions favouring emission will have disappeared before a stable arc is formed. On the other hand, the reactance reduces the value of current needed for stability and may, in that way, increase the danger of arc-back. It is probably unwise to assume that two different conditions giving the same value of F will have the same arc-back probability.

The chance of arc formation will vary also with the rate of rise of the inverse voltage on the anode. The factor F assumes the rate of rise to be infinite; this is acceptable for rough guidance, but the actual rate of rise should be considered in any careful analysis of the conditions of arc-back.

4. Experiments on effects of electrode contamination

4.1 *Insulating particles*

In experiments on an ignition system for aircraft engines, a 5000 pF capacitor was charged by the magneto to a voltage of a few kilovolts and then discharged, via a spark gap, through the primary of a transformer, the secondary of which was connected to a sparking plug. It was desirable for the gap, which was filled with hydrogen at a pressure of some hundreds of millimetres of mercury, to have a constant breakdown voltage; also, because of the limitation on the amount of energy available from the magneto, it was necessary to reduce to a minimum the energy losses in all components, including the spark gap.

The early form of the gap is shown in Fig. 1(*a*): the electrodes were sheet-metal pressings spaced apart by a cylindrical glass or ceramic spacer. Constancy of breakdown voltage was ensured by pre-ionization due to corona discharge at the junction of the spacer and the metal pressing. Loose insulating powder (magnesia or alumina) in later gaps shown in Fig. 1(*b*) produced a similar effect in stabilizing the breakdown voltage: the particles became charged with ions and the resulting field concentrations provoked electron emission. The initiation of emission at the insulator/conductor junction did not ensure the transition to an arc with a low voltage drop; this occurred regularly only when insulating particles were present on the cathode surface. The current in the gap and the voltage across it during the first half-cycle of current, in an equivalent circuit without the sparking-plug breakdown, is shown in Fig. 2 for three typical examples.

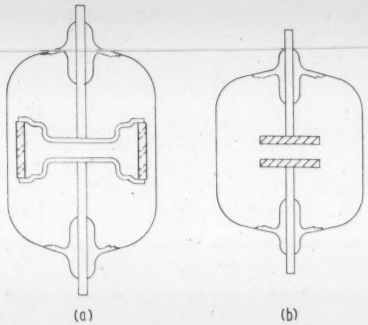


Fig. 1. Experimental, hydrogen-filled spark gaps. Glass bulb diameter 1¼ in.
(*a*) Electrodes spaced by glass or ceramic cylindrical member.
(*b*) No spacer..

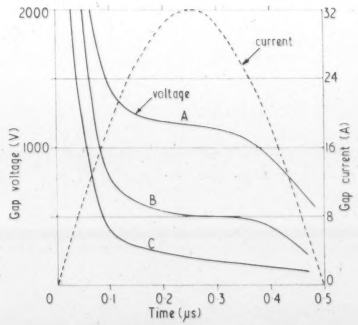


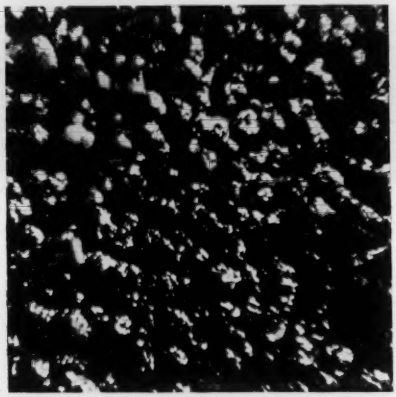
Fig. 2. Gap voltage and current in spark gaps as in Fig. 1(*a*). Curve A, nickel electrodes: I peak = 29·8 A; curve B, nickel electrodes: I peak = 31·3 A; curve C, aluminium electrodes: I peak = 33·5 A.

In many of the experiments the gap impedance was judged by the amount of damping observed on the oscillatory wave form of the open-circuit output voltage.

With early gaps having electrodes of sheet metal, the behaviour as regards damping was very erratic. Tests were made in a bell jar with the electrodes held 3½ mm apart, in air, by means of a spacer as shown in Fig. 1(*a*); the air pressure was reduced until the breakdown voltage was about 5 kV. The first experiments were made with dissimilar electrodes, one of nickel sheet and one of aluminium. The

breakdown voltage with the aluminium cathode was 10% lower than with the nickel cathode. In sparking tests, at about 300 sparks per second, there was at first no damping for either polarity; after three minutes there was some degree of damping with aluminium as cathode and critical damping with nickel. After about ten minutes the damping was less severe, and shortly afterwards it disappeared for both polarities.

The initial favourable results were due to the oxides adhering to the surface of each electrode. With continued sparking the oxide particles were removed from both electrodes by sputter, the effect being more pronounced at the nickel surface. The aluminium surface, however, subsequently became oxidized by the sparking in air, and in time this more than restored the position favourable to arc formation: under the microscope the aluminium appeared as if covered with metal globules, each having a white crystalline cap, presumably of aluminium oxide, as seen in the microphotograph of Fig. 3. The nickel electrode



0 0.1 mm

Fig. 3. Microphotograph of aluminium electrode after sparking in air in a gap similar to Fig. 1(*a*).

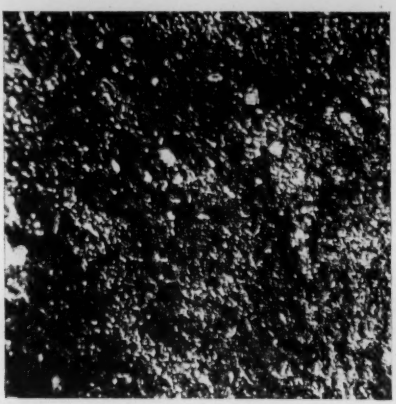
recovered a condition favourable to the formation of a low-impedance discharge when it became partially covered by patches of crystalline grains, over about 5% of its surface; these grains, which can be seen in the photograph of Fig. 4, are probably aluminium oxide sputtered on to the nickel from the other electrode.

In further experiments, the electrodes were of aluminium which had been oxidized, in some cases by sparking in air and in others by anodizing; in each case no damping was observed when the electrodes were used in hydrogen-filled spark gaps for about 30 to 50 hours, after which damping set in, the oxidized surface having disappeared because of sputter.

In spark gaps for service use, the continuous supply of insulating particles was ensured by forming the electrodes of a sintered mixture of metal and insulating powders (Knight and Welch 1944); a mixture of nickel with 7% by weight of alumina gave a satisfactory performance and life.

4.2 *Mercury contamination*

In the operation of ignitrons at 15 kv or more, it is important to ensure that free mercury is removed from the insulating bushing supporting the anode and also from the stress shields and electrode parts in their neighbourhood, in order to prevent voltage breakdown. In the arc-back tests described below, it was observed that the tendency to arc-back was



0 0.25 mm

Fig. 4. Microphotograph of nickel electrode after sparking in air in a gap similar to Fig. 1(*a*) with opposite electrode of aluminium.

increased by the presence of mercury on the anode itself. Mercury on an electrode increases the probability of breakdown by maintaining a local vapour pressure higher than that corresponding to the cold-spot temperature; there appears, however, to be some additional effect to account for arc formation at a surface contaminated by mercury.

Trials were made in experimental glass tubes, in mercury vapour, to study the formation of arc spots on different forms of cold cathode. The anode and cathode under test were connected to a capacitor-discharge circuit. Discharge was initiated by establishing ionization in the path, by passing a pulse of current, of about 0·5 A peak value, between an auxiliary anode and a hot cathode, in a separate circuit. The current flow thus initiated in the main path may be either an arc, with a visible cathode spot, or a high-impedance discharge.

Three forms of cathode were used: (*a*) a tungsten rod, 0·080 in. in diameter, sheathed with glass to within $\frac{1}{8}$ in. of the tip; (*b*) a nickel-sheet cylinder, $\frac{1}{4}$ in. long × $\frac{1}{4}$ in. diameter; (*c*) a graphite rod, $\frac{1}{2}$ in. long × $\frac{1}{2}$ in. diameter.

Table 1 gives the minimum capacitor voltage and peak current at which an arc occurred. The figures refer to the particular triggering method used: other conditions may give different results; for example, the voltage required is lower if the cathode is used as the anode of the priming discharge.

Table 1. *Conditions for arc formation at experimental cathodes*

Test No.	*Test cathode*	*Min. conditions for arc*
1	Glass-beaded tungsten	1000 V, 50 A
2	Nickel: surface wetted with Hg	300 V, 10 A
3	Graphite: surface wetted with Hg	500 V, 25 A

With unwetted nickel or graphite, there was no arcing at up to 2500 v, whether or not free mercury drops were present on the surface. For test No. 2, mercury was condensed on to the electrode by cooling with solid carbon dioxide applied to a boss fitted to the external terminal. For test No. 3, the graphite was first contaminated with nickel sputtered from an adjacent electrode. With the tungsten electrode the cathode spot formed at the glass/metal junction. On wetted surfaces the arc spots formed at random points, but generally only where the mercury was in a thin layer; also the conditions changed as the tests continued, requiring increasing values of voltage for arc formation.

The arc drop, measured with the cathode of test No. 2, varied linearly with current, from 28 v at 10 A to 80 v at 62 A. The vapour pressure was about 3 to 10 microns Hg. After many repeated pulses the arc drop tended to rise and then the arc failed (except at increased voltages). After a period allowed for cooling, arcing would again be obtained as before; this condition could be maintained for a prolonged period if the electrode was kept cool with solid carbon doxide (as described above).

The requirements for arc formation are lower the more effective the wetting by the mercury. If the nickel (in test No. 2) was tarnished before the mercury was deposited, the wetting was satisfactory in appearance, but the minimum values for arcing were 2000 v, 110 A, i.e. much greater than when the nickel was clean and the wetting more effective. A similar result is observed when the electrode becomes warm, a condition tending to reduce the degree of wetting.

Figure 5 shows a form of arc-transfer valve, in which a cathode spot is formed at an insulator/conductor junction

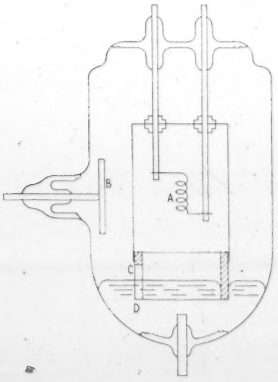


Fig. 5. Experimental arc-transfer valve with mercury-pool cathode and auxiliary hot cathode. *A*, auxiliary hot cathode; *B*, anode; *C*, slot in ceramic member; *D*, mercury-pool cathode.

on a cold cathode. In this case the main arc path is between an anode and a mercury-pool cathode. An auxiliary thermionic cathode is situated in an enclosed chamber, the wall of which, at its lower end, is of ceramic material, and dips into the mercury pool. The enclosed space communicates with the main chamber through a slot at mercury level.

One method of starting conduction is to initiate a priming discharge (for example with grid control) between the anode and the hot cathode, the latter being connected to the pool cathode through a current-limiting impedance. The priming discharge passes through the slot, close to the mercury surface. If the anode voltage is high enough an arc spot now forms at the ceramic/mercury junction, and the arc is transferred to the direct path from the anode to the pool cathode. The hot cathode need only carry a current of a fraction of an ampere, even though the main arc current may be thousands of amperes.

The arc spot forms by the same process that causes arc-back at insulating patches on rectifier anodes. But when the ceramic near the slot becomes contaminated with sputtered material and then wetted by mercury, the voltage required for arc formation and the current for arc stability are both lower than when the ceramic is clean.

Mercury contamination may be a contributory cause of arc-back in rectifiers. Mercury-wet patches are sometimes present on graphite anodes after a period of service; the mercury forms an amalgam with metal evaporated or sputtered on to the anode from walls or baffles by the action of arc spots or as a result of local heating due to the recombination of charged particles from the plasma.

Wetting by mercury favours the development and the stability of the arc; this is shown by the fact that when a cathode spot is anchored on a wetted metal surface projecting from a mercury pool, the arc drop is lower than when the spot is free, and the arc is stable at a lower current. According to the theory of von Bertele (1955), the work function of the mercury/metal alloy formed at the surface is lower than that of either member by itself. This would have no effect on arc formation when the mercury adheres in a thick layer; but in the present tests the likelihood of arcing increased as the electrode became warm and the excess mercury disappeared.

5. Arc-back tests

For a six-anode rectifier to have an arc-back probability not greater than one per year, each anode must have a probability of not more than one in six years. To obtain data useful in an assessment of rectifier performance, some kind of accelerated test is unavoidable. The rectifier circuit may be overloaded, or alternatively, if the power involved is considerable, test circuits may be used which apply simultaneously from low-power sources the voltage and current conditions leading to failure (Dobke 1951). Both methods are 'synthetic', and the conditions obtaining in any accelerated test may differ significantly from those in normal service, in respect of local temperatures, sputter and the migration of contaminating particles.

The occurrence of more frequent arc-back in the test circuit will itself introduce an artificial condition. When arc-back occurs it may remove its cause, the contamination being evaporated or removed bodily to some other part of the device; this causes ageing, reducing the probability of further arc-back. The effect may, however, be quite the reverse. Erosion of the anode may expose insulating particles hitherto hidden below the surface. Also, if an arc-back does not clear in a half-cycle but continues as an a.c. arc for one or more full cycles, the cathode of the next forward half-cycle may be on the metal envelope or on a baffle, instead of on the mercury pool; this may result in the evaporation or sputter of material which, settling as a contamination on the anode, increases the probability of further failure; this may be why arc-backs often occur in groups of two or more in quick succession.

A series of 'synthetic-circuit' tests, similar to those described by Kingdon and Lawton, were made on the type BK 56

ignitron (a water-cooled, steel-envelope, pentode-type ignitron rated at 20 kv, 150 A). The test circuit is shown in Fig. 6.

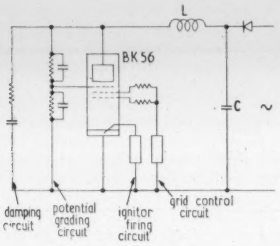


Fig. 6. Capacitor–discharge circuit for arc-back tests on high-voltage ignitron.

A capacitor C, charged to a variable voltage, was discharged 50 times a second through an inductance L by means of the ignitron under test. The ignitron current and inverse voltage are shown in Fig. 7. A damping circuit was connected

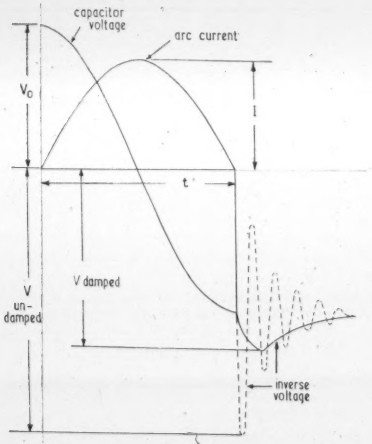


Fig. 7. Current and voltage oscillograms in circuit of Fig. 6 showing: current through ignitron, peak value I, duration t; capacitor voltage, initial value V_0; initial inverse voltage V across ignitron, with and without damping.

between the anode and the cathode to give partial suppression of any oscillatory overswing on the inverse voltage, except when high values of F were desired. The initial inverse voltage V, the peak current I and the current duration t were measured by means of an oscillograph and di/dt was calculated as equal to $\pi I/t$. Arc-backs were recorded by means of a magnetically operated counter. Tests, generally of several hours duration, were made under different conditions of arc-back factor, temperature and pre-treatment.

Typical results are illustrated in Fig. 8, in which the arc-back probability P is plotted against the factor F. (For a pulse frequency of 50 per second, one arc-back per hour represents a probability of $P = 1/50 \times 60 \times 60 = 5{\cdot}55 \times 10^{-6}$.) Before voltage was applied, all mercury was removed from the anode insulator bushings by preheating. The arc-back rate could still be reduced, however, by prolonging the heating until all mercury was evaporated also from the surface of the anode. Curve A records tests after careful ageing, and with an inlet-water temperature of 45° C. The line drawn through the points is the best-matching probability curve; it corresponds to $n = 14$, and gives a value of $k = 0{\cdot}0133$. Kingdon and Lawton's results with an 'aged' valve of similar anode/grid construction were matched to curves for $n = 12$, but having $k = 0{\cdot}0388$ for a cooling-water temperature of 60° C and $k = 0{\cdot}0194$ for 40° C. They took the 'aged' state as that at which 'the voltage to produce failure was not increasing rapidly with time'; the arc-back frequency in their tests was two orders greater, at comparable temperature, than in the 'aged' tests of curve A, Fig. 8.

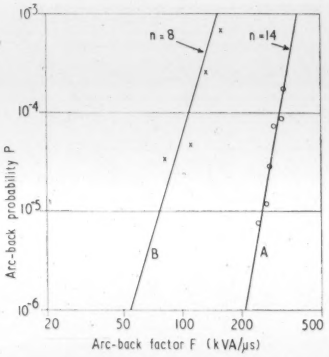


Fig. 8. Experimental results of tests on type BK 56 ignitron, showing arc-back probability as a function of arc-back factor F, matched to Poisson probability curves.

Curve A, optimum results after careful ageing; curve B, average over first hour after operation of the ignitron for one hour on continuous load, 100 A mean, 314 A peak.

Kingdon and Lawton observed ageing following arc-back and also deterioration when the anode was contaminated by dirt convected by the arc. Curve B, Fig. 8, shows the arc-back probability in the type BK 56 ignitron during the first hour of a test which was preceded by loading the valve for one hour with a rectified current of 100 A mean and 314 A peak value. These results are of greater significance than those of curve A in a forecast of the performance of the ignitron under practical conditions.

Arcing-back may lead either to ageing or to further deterioration. Curve B, Fig. 8, shows the average arc-back rate during the first hour after the load run. In prolonged tests the results conformed to the general pattern illustrated in Fig. 9; the arc-back rate increased steadily during the first hour and then slowly decreased, settling down to a fairly steady rate in about three hours, but at a level much higher than for an aged valve. The same can be seen in a series of tests, each of which lasted from five to ten minutes, and which were made in sequence as shown in Table 2.

Table 2. *Effect of arc-back rate on subsequent arc-back probability*

Factor F $VA/(\mu s)$	*Arc-backs per minute in successive tests*									
124	0									
162					4					
187		0				13		2		0
204			3						3	
220				14						0

In this case deterioration follows a high arc-back rate while ageing follows a low rate. In spite of this, a third observation is that a short 'burst' of arc-back at a high factor F is one of the best of the ageing agencies. In one case, with $F = 112$, the arc-backs were fairly evenly distributed at about twenty per hour during a $2\frac{1}{2}$ hour test; when the factor was increased, for two seconds only, to $F = 220$, with an inverse voltage of 22 kv, a continuation of the test at $F = 112$ showed only two arc-backs in one hour.

Curve B of Fig. 8 may be used in extrapolation to estimate the performance in rectifier service. The curve drawn through the experimental points is the probability curve for

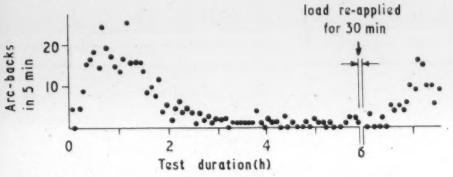


Fig. 9. Variation in rate of arc-back in type BK 56 ignitron during continuous tests in the circuit of Fig. 6, after operation for one hour on continuous load: note pattern of test results being repeated after re-application of load.

$n = 8$, assuming the constant k as determined by the 'aged' curve A. On this curve, $P = 10^{-10}$ occurs at $m = 0{\cdot}21$, corresponding to $F \simeq 16$. If the extrapolation is valid, then an average arc-back rate of one per six years of continuous service could be expected provided that F does not exceed 16 kVA/μs. However, the experimental points are obtained at values of F much higher than would be met in practice, and such an estimate involves a long extrapolation, which is generally undesirable.

Acknowledgments

The work was carried out in the Research Laboratory of A.E.I. (Rugby) Ltd. The author is indebted to colleagues for help: to L. Herbert for general collaboration; to A. E. Miller for the construction of apparatus and carrying out of many of the tests; and to D. F. Welch and A. Grieve who kindly supplied the curves of Fig. 2. Thanks are offered to the United Kingdom Atomic Energy Authority for permission to quote from the results of tests made in the course of an experimental programme under contract with them. The author also thanks the Directors of A.E.I. (Rugby) Ltd., for permission to publish the paper.

References

VON BERTELE, H., 1955, *Direct Curr.*, **2**, 143.

DÄLLENBACH, W., GERECKE, E., and STOLL, E., 1925, *Z. Phys.*, **26**, 10.

DOBKE, G., 1951, *A.E.G. Mitt.*, **41**, 165.

VON ISSENDORFF, J., 1929, *Elektrotech. Z.*, **50**, 1079.

VON ISSENDORFF, J., SCHENKEL, M., and SEELIGER, R., 1930, *Wiss. Veroff. Siemens*, **9**, 73.

KINGDON, K. H., and LAWTON, E. J., 1939, *Gen. Elect. Rev.*, **42**, 474.

KNIGHT, H. DE B., and WELCH, D. F., 1944, Brit. Pat. 580655.

LANGMUIR, I., 1928, *Z. Phys.*, **46**, 283.

MALTER, L., 1936, *Phys. Rev.*, **49**, 879.

SLEPIAN, J., and LUDWIG, L. R., 1932, *Trans Amer. Inst. Elect. Engrs*, **51**, 92.

STEINER, H. C., ZEHNER, J. L., and ZUVERS, H. E., 1944, *Trans Amer. Inst. Elect. Engrs*, **63**, 693.

Comprehensive study of the ion pumping of the noble gases

by B. COBIC, Institute of Nuclear Science, Vinca, P.O.B. 522, Belgrade, Yugoslavia, G. CARTER, B.Sc., Ph.D., Grad.Inst.P., and J. H. LECK, M.Eng., Ph.D., A.Inst.P., A.M.I.E.E., Department of Electrical Engineering, University of Liverpool

MS. received 8*th December* 1960, *in revised form* 25*th January* 1961

Abstract

The ion pumping of five noble gases by Bayard–Alpert ionization gauges has been studied extensively, employing an ultra-high-vacuum system which could be operated statically or dynamically. The dependence of the important pumping parameters, initial pumping speed and maximum quantity of gas pumpable, upon gas composition, electrode potentials and gauge temperature have been investigated. The relationship between instantaneous pumping speed and the quantity of gas pumped has been deduced, and observations upon the recovery of gas, following sorption, at ambient and elevated temperatures has also been made.

It is shown that the experimental results confirm a previously reported model of the sorption process, where gas ions enter a heterogeneous collection of capture sites of various energies of binding in the glass walls. The observed form of a static pump-down is interpreted in terms of the kinetic processes which can occur at these sites.

1. Introduction

CURRENT interest in ultra-high-vacuum systems has led to various general studies of the pumping of Bayard–Alpert ionization gauges (Bayard and Alpert 1950, Alpert 1953, Varnerin and Carmichael 1955, Young 1956, Bills and Carleton 1958). The present paper reports comprehensive measurements of the sorption of noble gases, chosen for their chemical inactivity, at glass surfaces. A static vacuum system has been chosen for this work as it has enabled previous measurements on a dynamic system (Carter and Leck 1959, 1961, Leck and Carter 1960) to be confirmed and has extended the scope of that work.

2. Experimental apparatus and technique

A schematic diagram of the glass vacuum system is shown in Fig. 1. The experimental Bayard–Alpert ionization gauge and a second monitoring gauge were evacuated by a well-trapped, two-stage mercury diffusion pump and could be sealed off by a conventional bakeable tap of the Alpert (1956) design. All the apparatus on the high-vacuum side of the diffusion pump could be heated to 350° c by means of a movable oven, and a second smaller oven served to heat the experimental gauge to 450° c. Gas could be admitted to the system at a controllable rate from an auxiliary reserve via a silicon carbide porous plug. The experimental gauge itself was a conventional Bayard–Alpert type with a nickel electron collecting grid, 2 cm diameter and 6 cm long, and a Pyrex glass envelope, 5 cm diameter and approximately 10 cm long.* This tube was mounted with its axis vertical. The pumping line and the electrical connections were all taken from the top of the tube, so that its whole length could easily be immersed in a Dewar vessel.

* The Pyrex brand glass was purchased from Jobling Ltd.

With this apparatus the following procedure was normally adopted to measure the pumping characteristics of the experimental gauge.

(1) When the background pressure had been reduced to approximately 10^{-9} torr as measured by the monitoring gauge (by pumping and baking to 350° c in the conventional manner) the experimental gauge was thoroughly degassed by baking to 450° c.

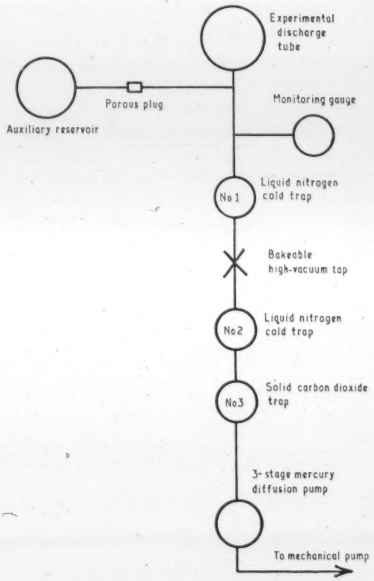


Fig. 1. Schematic diagram of the vacuum system.

(2) A controlled quantity of sample gas was admitted to the system and the tap closed so as to seal off the high-vacuum system at a pressure of the order of 10^{-5} to 10^{-4} torr. During this period the filament of the experimental gauge was operated at normal working temperature, but no potentials were applied to the other electrodes. Any small pumping action of the monitoring gauge (which operated with a low electron current so as to minimize its pumping action) and residual pumping through the valve were recorded by observing the small continuous pressure reduction.

(3) The experimental gauge potentials were then applied and the rapid pressure reduction which followed was observed.

All pressure measurements were made with the ionization gauge, which had a known calibration.

(4) When the pressure was reduced to about 10^{-7} torr, more gas was admitted from the auxiliary supply and the succeeding pump-down was observed. This could be repeated in order to pump as much gas as required.

Generally, as the walls of the gauge became charged with gas the pumping slowed and the rate of pressure reduction was observed to diminish. Typical pump-down cycles for argon are shown in Fig. 2. From the slope of these curves

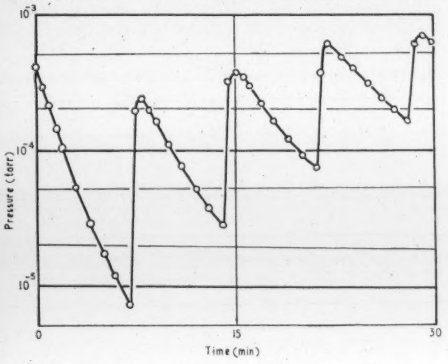


Fig. 2. Pump-down characteristics for argon. The pump operated continuously from time zero. The discontinuities in the curve occur when more gas is admitted to the system.

the speed of pumping of the gauge can be deduced, as the volume of the system is known to be 1·0 litre. Any gas released from the walls at room temperature could be recorded by removing the gauge potentials and recording the pressure increase. After a complete pump-down, gas was recovered by heating the gauge to 400° c. The trap No. 1, as indicated in Fig. 1, was refrigerated to −196° c for work with helium, neon, and argon, but, in order to prevent condensation, only to −78° c during experiments with krypton and xenon. This trapping ensured a low residual gas pressure throughout the work. The temperature of the glass of the experimental gauge could be controlled over a wide range by the oven (not more than a ±10 deg c spread over the length of the gauge at 300° c) or by immersion in a Dewar flask containing liquid nitrogen or a solid carbon dioxide acetone mixture.

3. Results

Pump-down characteristics, such as those shown in Fig. 2, have been obtained for all the five inert gases helium, neon, argon, krypton, and xenon. The measurements have been made for electron currents varied through the range 0·1 to 10 mA, electron accelerating voltages between 0 and 1000 v and gauge temperatures between −196 and +400° c. The values of speed that have been obtained from these measurements are set out in this section as functions of these parameters. In Fig. 3 the variation of the initial pumping speed S_0 is shown as a function of the electron emission current I_e for argon and operating with an applied electron accelerating potential of 250 volts. The small intercept on the speed axis is due to the combined effect of the pumping of the monitoring gauge and the leakage across the tap. In all subsequent results a correction has been made for this term. Fig. 4, also for argon, shows S_0 to be independent of gas pressure over a wide range. It is probably relevant that the break occurs at the point where the molecular mean free path becomes the same order of magnitude as the electron path length. In Fig. 5 the effective pumping speed S is shown as a

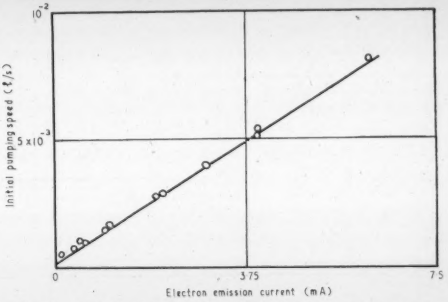


Fig. 3. The effective pumping speed of the experimental tube for argon as a function of the electron emission current.

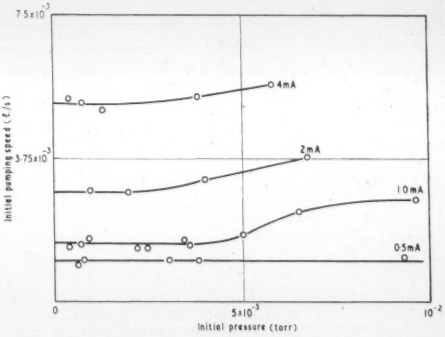


Fig. 4. Maximum pumping speed of the experimental tube S_0 plotted as a function of the gas pressure. (For argon with a grid-cathode accelerating voltage V_g of 250 v.)

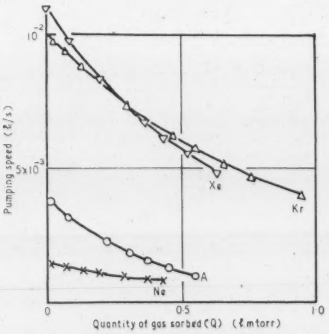


Fig. 5. Effective pumping speed as a function of the total quantity of gas pumped (V_g = 250 v).

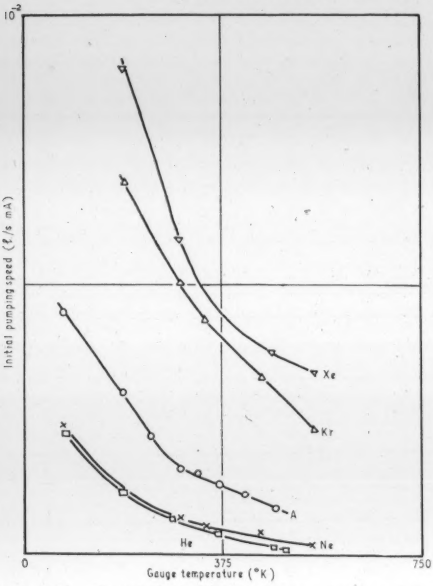


Fig. 6(*a*). The initial pumping speed S_0 as a function of the temperature of the glass envelope of the experimental tube.

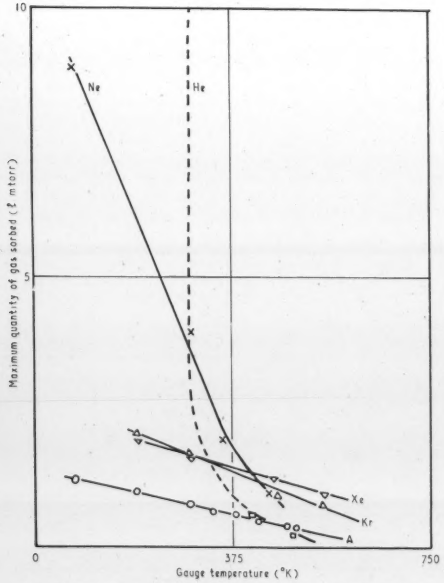


Fig. 6(*b*). The maximum quantity of gas sorbed Q_f as a function of the temperature of the glass envelope of the experimental tube.

function of Q, the total gas sorbed, for all the five inert gases. These results again were derived from curves such as those in Fig. 2 and S was calculated at the start of each pump-down, since at such points where the pressure was high any desorption was small compared to the pumping rate. In all the experiments, the saturation level Q_f, the value of sorption at which pumping ceases, was found to depend critically upon the electrode potentials of the discharge tube and the glass temperature, but to be quite independent of the electron current. This is illustrated in Figs 6 and 7 which show,

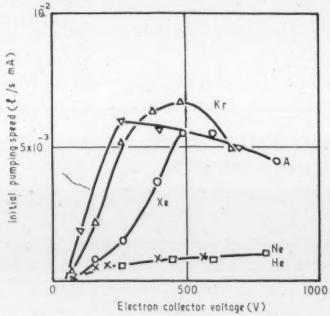


Fig. 7(*a*). The initial pumping speed S_0 as a function of the grid–cathode voltage V_g.

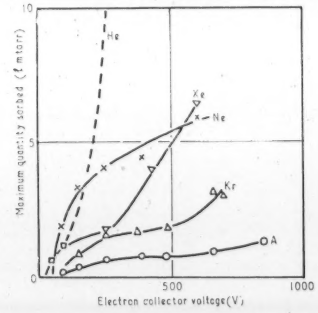


Fig. 7(*b*). The maximum quantity of gas pumped Q_f as a function of the grid–cathode voltage V_g.

respectively, S_0 and Q_f as functions of accelerating potential and tube temperature. Both S_0 and Q_f are tabulated in the Table for the five gases for the normal operating conditions of the Bayard–Alpert gauge.

The initial pumping speed S_0 and the maximum quantity of gas pumpable Q_f by the gauge operating at room temperature ($V_g = 250$ v)

Gas	S_0(*l./s mA* × 10^3)	Q_f(*l. mtorr*)
He	0·62	—
Ne	0·70	4·0
A	1·70	0·8
Kr	5·20	1·9
Xe	6·05	1·7

Since the glass is an insulator it is not possible to determine directly the exact ion flux to the wall. However, in an

auxiliary experiment (reported in previous work (Carter and Leck 1959, Leck and Carter 1960)) a metal shield was placed over the glass surface and the ion current to this plate measured for normal operating conditions. In this work the ion current to the wall can therefore be calculated for any operating conditions. It has been shown (Schwarz 1940, Carter 1959) that the walls generally stabilize at cathode potential, and under these circumstances it appears reasonable to assume that the ratio of the ion currents to the walls and collector remains constant. Thus from the measured pumping speed a simple calculation serves to give the ratio of the ions sorbed to the ion bombardment, i.e. the sticking factor K.

In Fig. 8 both Q_f and the maximum value of sticking

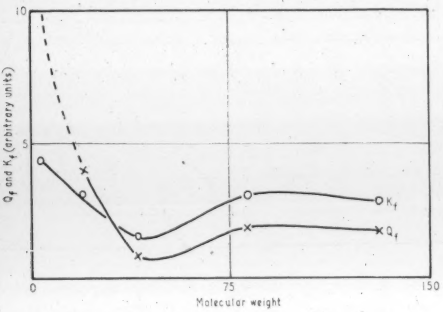


Fig. 8. The maximum sticking efficiency K_f and quantity of gas pumped Q_f as a function of the gas molecular weight.

factor K_f (i.e. at the commencement of sorption) are shown as functions of the molecular weight of the gas pumped; the similarity in characteristic is notable. Sticking efficiency is plotted against the electron accelerating voltage in Fig. 9.

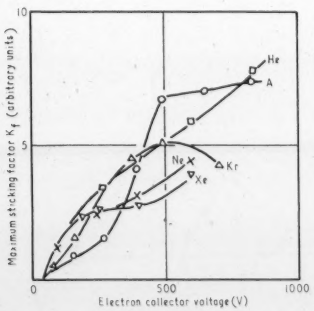


Fig. 9. The maximum sticking efficiency as a function of the voltage V_g.

The desorption following pump-down, either at room temperature or upon heating the gauge, was measured by observing the rise of pressure in the sealed system with the experimental gauge switched off. The quantity desorbed at room temperature was found to be of the order of 2% of the total sorbed gas. Comparison of the quantity of gas released with the quantity previously pumped, shows that heating to 350° C in the case of argon results in a 95% recovery, substantiating the earlier work. In the cases of krypton and xenon, since traps were refrigerated to only −78° C, considerable background gas evolved upon heating. However, it is believed that most of the gas pumped during these experiments was also recovered by heating to 350° C. An indication of the complete recovery was obtained by observing the initial speed in a subsequent pump-down; as this was substantially the same as in the previous experiment, it is reasonable to suppose that no gas remained in the walls.

4. Discussion

Earlier work (Carter and Leck 1961) has shown that the probable process of sorption is one in which gas ions are impelled beneath the glass surface into favoured capture sites, which vary in size because of varying interatomic spacings in the glass. This leads to a spectrum of energies of binding between the sorbed gas and the glass. These energies appear to be of similar magnitude to energies of activation for diffusion of inert gases in glass, and to range from 25 to 50 kcal/mole.

In the earlier work it was found that the trapped molecules could be released either thermally or by a further ion bombardment. The present results support this model, and it becomes possible to describe the pump-down in terms of two main competing processes. The first is the trapping of ions in favoured sites, and the second is the evolution of trapped gas by thermal desorption or by sputtering, of either the gas directly or of the surrounding glass, in the continuing discharge. (Evolution of trapped gas by sputtering was first reported by Schwarz (1944) and is the subject of further investigation in these laboratories.)

Initially—when the total sorption is very small—the evolution process will be unimportant and the factor K indicates the true sticking efficiency of the ions. Subsequently, as the sorption increases, the evolution becomes increasingly more important and the net speed and effective value of K naturally become a complex function of the sorption; therefore, curves such as those in Figs 2 and 5 may be anticipated. When the discharge is stopped thermal evolution will be the only process operating. As more gas is pumped, the recovery processes must become increasingly important until eventually they balance the pumping, thus giving a dynamic equilibrium. The quantities S_0 and Q_f are thus representative of the sticking and a combination of sticking and evolution efficiencies, and are both linearly dependent upon the number of sites available for sorption.

The quantities S_0 and Q_f fall to zero at approximately 400° C for argon (the fact that both fall to zero at the same temperature indicates that the sticking efficiency, which conditions both, is really zero at this temperature) showing that there are no sites having an energy of binding above 50 kcal/mole, again in agreement with previous work. The fact that this limiting temperature is lower for neon and higher for both krypton and xenon is again in agreement with earlier observations, and to be expected from the relative size of the gas atoms. The increase in S_0 and Q_f at lower than room temperature indicates the existence of further sites of energy of activation considerably below 25 kcal/mole, sites which cannot contribute significantly to the pumping at room temperature because of rapid thermal desorption at this temperature. Some evidence for the existence of these sites can be seen in the recovery of gas which occurs at room temperature when the discharge is switched off. The calculated energy of these sites is 24 kcal/mole, but the d'Arrhenius plot shows curvature, indicating a band of energies around this figure. The number

of these sites is relatively small—of the order of a few per cent—and they are therefore filled comparatively quickly during sorption. (The desorption at room temperature has been shown by a number of workers to be reciprocal time-dependent (Varnerin and Carmichael 1957, Fox and Knoll 1961). From this it has been inferred that a diffusion mechanism has a controlling influence on the desorption. This evidence does not in any way conflict with the hypothesis put forward above. In the present work the major fraction of the desorption takes place above 100° c, where a diffusion mechanism controlling at room temperature would be too fast to be observed.)

It is clear from Fig. 8 that the characteristic forms of the sticking factor K_f and Q_f versus atomic weight curves are very similar. Since both are linear functions of the site availability, the similarity may be anticipated. These K_f graphs in Fig. 8 show that both K_f and Q_f are large for helium and then decrease with atomic weight, exhibiting a minimum for argon before slowly increasing. The initial behaviour may be anticipated, since helium and neon are both relatively small atoms and can be expected to penetrate the glass relatively easily. The subsequent small increase with increasing atomic weight is a little unexpected and cannot be explained in simple terms. It may, for example, be due to the variation in site availability with atomic diameter, or even to a dependence upon the multiply charged ion formation in krypton and xenon, which lead to higher bombarding energies.

The observed increase of K_f and Q_f with electron accelerating potential may also be anticipated, since the more energetic ions can be expected to penetrate to sites forbidden to lower energy ions, resulting in an overall increase in the site availability. Since the electron and ion trajectories in the Bayard–Alpert gauge are complex and the ion bombardment of the walls is inadequately known, it is not proposed to offer an analysis of the form of the increase of K_f and Q_f with voltage.

In many respects the behaviour of helium is anomalous. Very large quantities of this gas can be pumped and only a relatively small amount recovered by thermal desorption. The quantities S_0 and Q_f were found to depend to some extent upon the gas pressure and electron emission current. This can be interpreted as indicating that when helium is sorbed the molecules diffuse further into the glass, setting up a concentration gradient allowing continuous sorption. Equilibrium is obtained when this diffusion balances the net sorption at the surface and thus depends upon the bombardment rate. The anomalous behaviour for helium, reported by Varnerin and Carmichael (1955) and Young (1956), that the sorption depends upon a metal surface layer, has not been noticed in the present work. The results are, however, not incompatible if it is assumed that here equilibrium surface conditions had been reached before the first observations were made. Thus the surface barrier prevents diffusion out of, but not into, the interior of the glass.

5. Conclusions

New information concerning the performance of Bayard–Alpert gauges as vacuum pumps has been presented in this paper. The characteristic parameters of the gauge pump, initial pumping speed and maximum quantity pumpable, under various conditions of tube temperature, electrode potentials and emission current for five of the noble gases have been comprehensively studied and are given here. The reproducibility of the results and the complete agreement with the previous experiments indicates the reliability of both the present and the earlier work. The conclusions derived from this work are in agreement with the proposals put forward earlier concerning the mechanism of retention of pumped gas in the glass, i.e. the capture sites of varying energies of binding, and has led to a modification of the model to account for the apparently unlimited sorption of helium, in which a helium gas atom is initially trapped in a capture site but can subsequently diffuse into the glass.

Observations on the thermal desorption of the gases from the glass subsequent to sorption also agree with the proposed sorption model, and indicate that, in the case of these gases, desorption can be expressed in terms of gas evolution from a spectrum of capture sites. The room temperature recovery has been found to be only a small percentage of the total quantity pumpable, but under static conditions contributes to the ultimate pressure attainable in the system. It appears that most efficient pumping of noble gases is attained by operating at high electron emission currents, high electron accelerating energies and low tube temperatures.

Acknowledgments

The authors would like to thank Professor Craggs for his advice and the continuous interest he has shown in this work. One of us (B. C.) would like to thank the Institute for Nuclear Sciences, Belgrade, for financial assistance, and the authorities of Liverpool University who provided the facilities for her work at Liverpool.

References

ALPERT, D., 1953, *J. Appl. Phys.*, **24**, 860.
—— 1956, *Rev. Sci. Instrum.*, **22**, 536.
BAYARD, R. T., and ALPERT, D., 1950, *Rev. Sci. Instrum.*, **21**, 571.
BILLS, D. G., and CARLETON, N. P., 1958, *J. Appl. Phys.*, **29**, 692.
CARTER, G., 1959, *Nature, Lond.*, **183**, 1619.
CARTER, G., and LECK, J. H., 1959, *Brit. J. Appl. Phys.*, **10**, 364.
—— 1961, *Amer. Vac. Society Symposium*, 1960, in the press.
FOX, R. E., and KNOLL, J. S., 1961, *Amer. Vac. Society Symposium*, 1960, in the press.
LECK, J. H., and CARTER, G., 1960, *Advances in Vacuum Science and Technology*, Vol. I, 463 (London: Pergamon Press Ltd.).
SCHWARZ, H., 1940, *Z. Phys.*, **117**, 23.
—— 1944, *Z. Phys.*, **122**, 437.
VARNERIN, L. J., and CARMICHAEL, J. H., 1955, *J. Appl. Phys.*, **26**, 782.
—— 1957, *J. Appl. Phys.*, **28**, 913.
YOUNG, J. R., 1956, *J. Appl. Phys.*, **27**, 926.

End effects in a falling-sphere viscometer

by A. D. MAUDE, B.Sc., Ph.D., A.Inst.P., University College of Wales, Aberystwyth

MS. received 11*th January* 1961

Abstract

The calculations of Stimson and Jeffery of the force on two spheres moving coaxially through a very viscous liquid are extended to the case when the spheres move at different rates. This result is used to calculate the effect of the bottom of the vessel in a falling sphere viscometer and that of the upper surface of the liquid. Approximate solutions are also found in a form suitable for practical use.

Introduction

When the viscosity of a liquid is measured by means of a falling-sphere viscometer, corrections have to be made to Stokes's famous equation to allow for the finite size of the vessel containing the liquid. If the vessel consists of a long cylinder, and if the sphere falls along the axis of it an equation for the effect of the walls has recently been given by Bohlin (1960) which is accurate to the tenth power of the ratio of the sphere to the cylinder diameters. The effect of the upper surface and of the closed lower end of the vessel is, however, still somewhat obscure.

Lorentz (1906–7) has deduced an approximate equation to give the effect of a sphere falling towards an infinite rigid plane, which Hopper and Grant (1948) found to be valid only when the sphere was more than twenty times its radius from the plane. The author has not found any work on the effect of the free upper surface of the liquid.

An exact solution for the creeping flow of a liquid around a sphere as it approaches an infinite rigid plane has been found together with a similar solution for the flow around a sphere as it approaches an infinite plane surface on which there is no traction force. This latter is an approximation for the flow as a small rigid sphere approaches (or leaves) a liquid surface, as long as movement of the sphere does not appreciably alter the shape of the surface.

The interaction of the effects of the wall and the ends of the vessel, which may often be important, has not been investigated.

Theoretical

Stimson and Jeffery (1926) have given a theoretical solution to the motion of two solid spheres moving with equal constant velocities parallel to their line of centres through a viscous fluid at such a rate that the inertial terms of the equation motion of the fluid may be neglected. The method used is immediately applicable to the similar problem in which sphere velocities differ. It was found most convenient to solve the problem for the special case where the spheres move apart with equal and opposite velocities, and to obtain the solution for any other pair of velocities by combining a solution of this form with one of the form derived by Stimson and Jeffery.

The reader may consult the original paper for the details of the calculation, so that only the result is quoted here.

The flow is described in terms of Stokes's stream function, so that the velocity components of the fluid u and v are

$$u = \frac{1}{\omega}\frac{\delta\psi}{\delta z},\ v = -\frac{1}{\omega}\frac{\delta\psi}{\delta\omega}$$

where ω and z are cylindrical coordinates. Bipolar coordinates ξ and η are used, such that

$$\omega = \frac{a\sin\eta}{\cosh\xi - \cos\eta},\ z = \frac{a\sinh\xi}{\cosh\xi - \cos\eta}.$$

The spheres are chosen to be $\xi = \alpha$ and $\xi = \beta$, ($\alpha \geqslant 0$, $\beta \leqslant 0$) and the sphere radii r_1, r_2 and the distances of their centres from (and on opposite sides of) the origin l_1, l_2 are given by $r_1 = a\,\mathrm{cosech}\,\alpha$, $r_2 = -a\,\mathrm{cosech}\,\beta$, $l_1 = a\coth\alpha$, $l_2 = -a\coth\beta$. The solution for the perturbation velocity is then shown to be

$$\psi = (\cosh\xi - \cos\eta)^{-3/2}\sum_{n=1}^{\infty} U_n V_n$$

where, with the usual Legendre function notation,

$$V_n = \mathrm{P}_{n-1}(\cos\eta) - \mathrm{P}_{n+1}(\cos\eta)$$

and

$$U_n = A_n\cosh(n-\tfrac{1}{2})\xi + B_n\sinh(n-\tfrac{1}{2})\xi + C_n\cosh(n+\tfrac{3}{2})\xi + D_n\sinh(n+\tfrac{3}{2})\xi$$

where A_n, B_n, C_n, D_n are found from the boundary conditions on the spheres.

It is shown that the forces on the spheres are

$$F_1 = -\frac{\mu\pi 2\sqrt{2}}{a}\sum_{n=1}^{\infty}(2n+1)(A_n - B_n + C_n - D_n)$$

for the sphere $\xi = \alpha$, and

$$F_2 = -\frac{\mu\pi 2\sqrt{2}}{a}\sum_{n=1}^{\infty}(2n+1)(A_n + B_n + C_n + D_n)$$

for the sphere $\xi = \beta$, where μ is the fluid viscosity.

Using the boundary conditions $u = 0$, $v = V$ on the surface of both spheres Stimson and Jeffery obtain

$$\begin{aligned}\Delta A_n = {} & (2n+3)K[4\exp\{-(n+\tfrac{1}{2})(\alpha-\beta)\}\sinh(n+\tfrac{1}{2})(\alpha-\beta) \\ & + (2n+1)^2\exp(\alpha-\beta)\sinh(\alpha-\beta) \\ & + 2(2n-1)\sinh(n+\tfrac{1}{2})(\alpha-\beta)\cosh(n+\tfrac{1}{2})(\alpha+\beta) \\ & - 2(2n+1)\sinh(n+\tfrac{3}{2})(\alpha-\beta)\cosh(n-\tfrac{1}{2})(\alpha+\beta) \\ & - (2n+1)(2n-1)\sinh(\alpha-\beta)\cosh(\alpha+\beta)]\end{aligned}$$

$$\begin{aligned}\Delta B_n = {} & -(2n+3)K[2(2n-1)\sinh(n+\tfrac{1}{2})(\alpha-\beta)\sinh(n+\tfrac{1}{2})(\alpha+\beta) \\ & - 2(2n+1)\sinh(n+\tfrac{3}{2})(\alpha-\beta)\sinh(n-\tfrac{1}{2})(\alpha+\beta) \\ & + (2n+1)(2n-1)\sinh(\alpha-\beta)\sinh(\alpha+\beta)]\end{aligned}$$

$$\Delta C_n = -(2n-1)K[4\exp\{-(n+\tfrac{1}{2})(\alpha-\beta)\}\sinh(n+\tfrac{1}{2})(\alpha-\beta) - (2n+1)^2\exp\{-(\alpha-\beta)\}\sinh(\alpha-\beta) + 2(2n+1)\sinh(n-\tfrac{1}{2})(\alpha-\beta)\cosh(n+\tfrac{3}{2})(\alpha+\beta) - 2(2n+3)\sinh(n+\tfrac{1}{2})(\alpha-\beta)\cosh(n+\tfrac{1}{2})(\alpha+\beta) + (2n+1)(2n+3)\sinh(\alpha-\beta)\cosh(\alpha+\beta)]$$

$$\Delta D_n = (2n-1)K[2(2n+1)\sinh(n-\tfrac{1}{2})(\alpha-\beta)\sinh(n+\tfrac{3}{2})(\alpha+\beta) - 2(2n+3)\sinh(n+\tfrac{1}{2})(\alpha-\beta)\sinh(n+\tfrac{1}{2})(\alpha+\beta) + (2n+1)(2n+3)\sinh(\alpha-\beta)\sinh(\alpha+\beta)]$$

where $\Delta = 4\sinh^2(n+\tfrac{1}{2})(\alpha-\beta) - (2n+1)^2\sinh^2(\alpha-\beta)$

and $$K = Va^2\frac{n(n+1)}{\sqrt{2(2n-1)(2n+1)(2n+3)}}.$$

Using the same method and the boundary conditions $u = 0$ $v = -V$ on the sphere $\xi = \alpha$, and $u = 0$ $v = +V$ on sphere $\xi = \beta$ the following expressions are obtained:

$$\Delta A'_n = +(2n+3)K[2(2n-1)\sinh(n+\tfrac{1}{2})(\alpha-\beta)\sinh(n+\tfrac{1}{2})(\alpha+\beta) - 2(2n+1)\sinh(n+\tfrac{3}{2})(\alpha-\beta)\sinh(n-\tfrac{1}{2})(\alpha+\beta) - (2n+1)(2n-1)\sinh(\alpha-\beta)\sinh(\alpha+\beta)]$$

$$\Delta B'_n = -(2n+3)K[-4\exp(n+\tfrac{1}{2})(\alpha-\beta)\sinh(n+\tfrac{1}{2})(\alpha-\beta) - (2n+1)^2\exp(\alpha-\beta)\sinh(\alpha-\beta) + 2(2n-1)\sinh(n+\tfrac{1}{2})(\alpha-\beta)\cosh(n+\tfrac{1}{2})(\alpha+\beta) - 2(2n+1)\sinh(n+\tfrac{3}{2})(\alpha-\beta)\cosh(n-\tfrac{1}{2})(\alpha+\beta) + (2n-1)(2n+1)\sinh(\alpha-\beta)\cosh(\alpha-\beta)]$$

$$\Delta C'_n = -(2n-1)K[+2(2n+1)\sinh(n-\tfrac{1}{2})(\alpha-\beta)\sinh(n+\tfrac{3}{2})(\alpha-\beta) - 2(2n+3)\sinh(n+\tfrac{1}{2})(\alpha-\beta)\sinh(n+\tfrac{1}{2})(\alpha-\beta) - (2n+1)(2n+3)\sinh(\alpha-\beta)\sinh(\alpha+\beta)]$$

$$\Delta D'_n = +(2n-1)K[-4\exp\{-(n+\tfrac{1}{2})(\alpha-\beta)\}\sinh(n+\tfrac{1}{2})(\alpha+\beta) + (2n+1)^2\exp\{-(\alpha-\beta)\}\sinh(\alpha-\beta) + 2(2n+1)\sinh(n-\tfrac{1}{2})(\alpha-\beta)\cosh(n+\tfrac{3}{2})(\alpha+\beta) - 2(2n+3)\sinh(n+\tfrac{1}{2})(\alpha-\beta)\cosh(n+\tfrac{1}{2})(\alpha+\beta) - (2n+1)(2n+3)\sinh(\alpha-\beta)\cosh(\alpha+\beta)].$$

It is now possible to obtain the forces on two spheres moving with different velocities along the line of centres in an infinite fluid. For example, if one sphere is stationary and a second approaches it with a velocity V, the force on the stationary sphere is

$$F_1 = -\frac{\mu\pi\sqrt{2}}{a}\sum_{n=1}^{\infty}(2n+1)(A_n+A'_n+B_n+B'_n+C_n+C'_n+D_n+D'_n)$$

and on the moving sphere is

$$F_2 = -\frac{\mu\pi\sqrt{2}}{a}\sum_{n=1}^{\infty}(2n+1)(A_n+A'_n-B_n-B'_n+C_n+C'_n-D_n-D'_n).$$

This case is of particular importance if the radius of the stationary sphere is very large, as this then represents the case of a sphere approaching an infinite rigid plane such as the bottom of the containing vessel of a falling sphere viscometer. In this case $\alpha = 0$, and the distance of the sphere centre from the plane l and the sphere radius r are related to β and a by the equations

$$\cosh\beta = +l/r,\ a^2 = l^2 - r^2,\ \beta \text{ negative.}$$

The force on the sphere then becomes

$$F_1 = 6\pi\mu Vr\lambda_1$$

$$\lambda_1 = -\tfrac{1}{4}\sinh\beta\sum_{n=1}^{\infty}\left\{\frac{n(n+1)\,[8e^{(2n+1)\beta} + 2(2n+3)(2n-1) - (2n+1)(2n-1)\,e^{2\beta} - (2n+1)(2n+3)\,e^{-2\beta}]}{(2n-1)(2n+3)\,[4\sinh^2(n+\tfrac{1}{2})\beta - (2n+1)^2\sinh^2\beta]}\right\}. \quad (1)$$

The force on the plane reduces to the same expression, with the opposite sign.

The boundary conditions pertaining to a sphere of radius r falling at a velocity V from below a free plane liquid surface are automatically satisfied by equal spheres separating, each with a velocity V (the free surface being the plane midway between them), as on this plane there will be no velocity perpendicular to, nor any stress parallel to it. Thus putting $\beta = -\alpha$, the sphere radius r, the distance of the sphere centre below the free surface d, and the parameter a are related by

$$\cosh\alpha = d/r,\ a^2 = d^2 - r^2$$

$$F = -\frac{\mu\pi\sqrt{2}}{a}\sum_{n=1}^{\infty}(2n+1)(A'_n+B'_n+C'_n+D'_n) = 6\pi\mu Vr\lambda_2$$

where

$$\lambda_2 = -\frac{4}{3}\sinh\alpha\sum_{n=1}^{\infty}\left[\frac{n(n+1)}{(2n-1)(2n+3)} \times\left\{1 - \frac{4\cosh^2(n+\tfrac{1}{2})\alpha + (2n+1)^2\sinh^2\alpha}{2\sinh(2n+1)\alpha - (2n+1)\sinh 2\alpha}\right\}\right]. \quad (2)$$

This may be compared with result when the spheres move in the same direction given by Stimson and Jeffery

$$F = 6\pi\mu Vr\lambda_3$$

$$\lambda_3 = +\frac{4}{3}\sinh\alpha\sum_{n=1}^{\infty}\left[\frac{n(n+1)}{(2n-1)(2n+3)} \times\left\{1 - \frac{4\sinh^2(n+\tfrac{1}{2})\alpha - (2n+1)^2\sinh^2\alpha}{2\sinh(2n+1)\alpha + (2n+1)\sinh 2\alpha}\right\}\right]. \quad (3)$$

In Stimson and Jeffery's paper 2/3 appears instead of 4/3, presumably due to a printing error.

Discussion

The force on a sphere approaching a rigid plane has been investigated experimentally by Hopper and Grant (1948), who used oil drops falling through air. As the viscosity of the oil was much greater than that of the air, and as the drops were small enough to remain spherical (a few microns diameter), they could be considered as solid spheres. These experimental results were summarized as follows:

$$\text{Force} = 6\pi\mu rV\lambda_1''$$

$$\lambda_1'' = 1 + K_1\frac{r}{l} + K_2\left(\frac{r}{l}\right)^2$$

$$K_1 = 1{\cdot}08 \pm 0{\cdot}05$$
$$K_2^* = 1{\cdot}4 \pm 1{\cdot}0.$$

If Eqn (1) is expanded as a power series in terms of r/l we get

$$\lambda_1 = 1 + \frac{9}{8}\frac{r}{l} + \left(\frac{9}{8}\right)^2\left(\frac{r}{l}\right)^2 + \ldots \quad (4)$$

$$= 1 + 1{\cdot}125\,\frac{r}{l} + 1{\cdot}266\left(\frac{r}{l}\right)^2 + \ldots$$

It will be seen that this is in good agreement with the results of Hopper and Grant. The result also agrees with the theoretical result given by Lorentz (1906–7)

$$\lambda_1 = 1 + \frac{9}{8}\frac{r}{l} \text{ for small } \frac{r}{l}.$$

The form of Eqn (4) suggests that we may write, as a useful approximation,

$$\lambda_1 = 1\Big/\left(1 - \frac{9}{8}\frac{r}{l}\right). \quad (5)$$

In the same way we find from Eqn (2)

$$\lambda_2 = 1 + \frac{3}{4}\frac{r}{d} + \left(\frac{3}{4}\right)^2\left(\frac{r}{d}\right)^2 \quad (6)$$

suggesting

$$\lambda_2 = 1\Big/\left(1 - \frac{3}{4}\frac{r}{d}\right). \quad (7)$$

These approximations are of course better for large values of l/r and d/r. The limit of their usefulness may be judged from the Table.

β	l/r	λ_1 *true*	λ_1 *approx.*
-2	3·762	1·412	1·426
$-1{\cdot}5$	2·352	1·838	1·917
-1	1·543	3·039	3·961

α	d/r	λ_2 *true*	λ_2 *approx.*
2	3·762	1·247	1·249
1·5	2·352	1·463	1·468
1	1·543	1·973	1·945

References

BOHLIN, T., 1960, *K. Tekn. Högsk. Handl.*, **155.**

HOPPER, V. D., and GRANT, A. M., 1948, *Aust. J. Sci. Res.*, A, **1,** 28.

LORENTZ, H. A., 1906–7, *Abhandlungen über theoretische Physik*, Vol. 1, 23 (Leipzig: Teubner).

STIMSON, M., and JEFFERY, G. B., 1926, *Proc. Roy. Soc.*, A, **111,** 110.

Technique for the rapid, accurate and strain-free machining of metallic single crystals

by M. COLE, M.A., Ph.D., I. A. BUCKLOW, M.A., Ph.D., Metals Research Ltd., 91 King Street, Cambridge, and C. W. B. GRIGSON, M.A., Ph.D., Department of Engineering, University of Cambridge

MS. received 28th November 1960

Abstract

Spark planing, a new spark erosion method for the production of flat smooth metal surfaces, is described and its application to single metal crystals and to metallographic preparation is illustrated. The use of conventional spark erosion methods for crystal cutting and forming is also described. The spark planing technique is far more rapid and accurate than chemical or electrochemical machining and it causes very much less damage to the surface than the most careful grinding operation. Spark planed surfaces have been examined by x-ray diffraction, reflection electron microscopy and a Talysurf tracer: these results are presented and the nature and extent of surface damage to various metals is discussed.

1. Introduction

THE preparation of unstrained metal single-crystal specimens to close dimensional and crystallographic tolerances has always been a difficult and tedious process. Mechanical methods such as turning and grinding, even when very light cuts are taken, always introduce considerable strain, especially in the surface layers; if the severely damaged layer is removed by electropolishing or etching, a relatively strain-free surface results, but the dimensional accuracy is lost, especially over large areas. The technique described here is capable of preparing metallic surfaces to a high degree of dimensional accuracy without producing a deep distorted layer.

The method, which consists of the removal of metal by a rapid series of sparks, was first reported by Lazarenko (1946) for the accurate machining of hard metal dies. James and Milner (1953) described a similar technique for the strain-free cutting of metal specimens, but the resultant surface was rough and further mechanical preparation was necessary to produce a flat and smooth finish. The present modification of the spark-machining technique can cut specimens to various shapes and also produce an accurate, flat 'satin finish' (10 μin.) surface which is, in most cases, virtually strain free.

2. Spark-machining principle

The principle of spark machining is illustrated in Fig. 1. The specimen to be machined S is immersed in a tank containing an insulating liquid of high dielectric constant and low viscosity, such as kerosene or transformer oil. The capacitor C is charged through the resistor R to the breakdown potential V of the dielectric in the work gap d; a spark discharge then takes place across the gap. Capacitor C is then recharged through R and the cycle repeated; the time constant CR must be large enough to prevent the formation of a continuous arc. The breakdown voltage V is proportional to the work gap d, which must be controlled within close limits for efficient operation; V is therefore used to control a servo-system which maintains d at its optimum value of approximately 0·0015 in. The accuracy and response time of the servo-system and its amplifier must be such as to ensure that the tool never hits the work, a feature essential for strain-free machining.

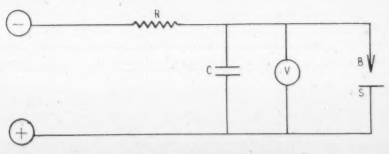


Fig. 1. Circuitry for spark machining.

3. Spark cutting

A specimen is cut through by maintaining the blade automatically at a distance from the specimen such that continual sparking occurs, thus producing a slot which is larger than the blade by the breakdown distance of the dielectric at the working voltage (about 0·0015 in.). If a high value of $\frac{1}{2}CV^2$ is used, cutting will be rapid and a rough surface will be produced; a low value will reduce the cutting rate, but will produce a smoother surface. With $V = 200$ volts, R = 12 ohms and C = 200 μF, a 1 in. diameter bar of aluminium can be cut in 5 minutes and a similar copper rod in 30 minutes. The process can also be used to cut specimens of various shapes in a required orientation from a larger block; thus Fig. 2 illustrates the use of a tubular cutter to cut out a cylinder or core a specimen. Cores of other than circular section can be trepanned by this method.

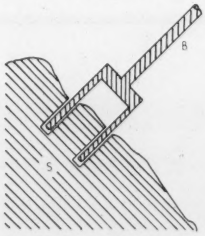


Fig. 2. Coring tool.

4. The production of flat, smooth surfaces

If the blade B of Fig. 1 is replaced by a flat horizontal surface, and if this is brought close to a rough specimen, sparking occurs between the high spots on the specimen and the cutter surface. The high spots are eventually eroded, but the metal removed from the specimen builds up as a suspension in the narrow gap between the electrodes and

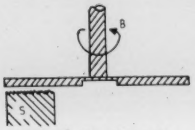


Fig. 3. Planing tool.

forms a bridge of conducting particles, which allows the spark to travel farther through the dielectric. These particles readily escape from the edges of the work but build up in the centre. The bridging effect is therefore pronounced in the centre; thus more metal is removed from this region and a concave surface results. This defect can be overcome by using a rotating disk (patent applied for) as the cutting tool; the disk should have a central well and radial slots and the specimens should be mounted near the edge of the disk B, as in Fig. 3. Rotation of the disk causes paraffin to flow radially outwards, thus ensuring a fresh supply over the whole surface. The specimen is initially machined with high energy sparks to produce a flat but rough surface, and then the energy of the sparks is successively reduced until a very flat, reasonably smooth surface results. Such a surface is suitable for metallographic examination after a light etch or very light electropolish.

5. The nature of spark-machined surfaces

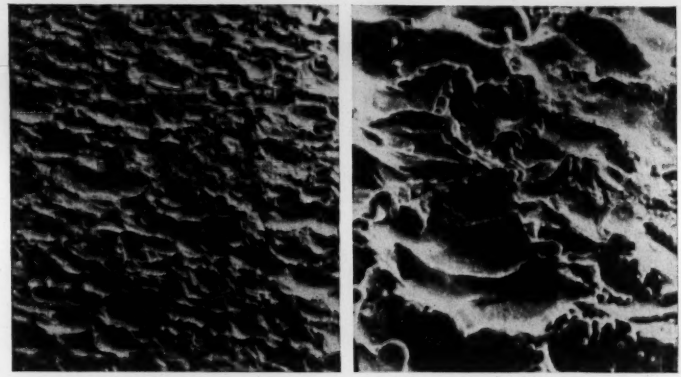
Spark-machined surfaces of many metals and other conducting substances have been examined by x-ray diffraction, reflection electron microscopy and a Talysurf tracer (Taylor, Taylor and Hobson).

Laue photographs from spark-planed single-crystal surfaces of many metals show sharply defined spots; in those few cases where there is surface damage or contamination, a light etch is sufficient to produce sharp reflections.

Reflection electron microscopy of spark-planed surfaces of the many metals and alloys so far investigated shows them to be very similar in appearance. In Fig. 4 are reflection micrographs* of typical surfaces produced by planing aluminium in a medium-light range. The craters are approximately 25 to 30 μ across and about 2 μ deep.

Talysurf measurements of the same specimens confirm these figures and give a centre line average surface texture figure of 10 μin. Chemical contamination in varying degrees can occur and consists of cathode material and carbon from decomposition of the paraffin. Contamination is slight and confined to a very thin surface layer; it can be minimized by suitable choice of cathode material and electrical parameters and is removed by a light etch.

The physical condition of spark-machined surfaces is being studied and will be the subject of another paper to be published shortly; however, preliminary results suggest that on soft cubic metals the damage on the fine ranges, if any, is confined to a layer of the order of a few microns deep. Zinc crystals frequently cleave along the basal plane if machined below 100° C, but do not appear to suffer any damage if machined above this temperature. Twinning to a depth of about 1 μ on α-uranium has been reported by Hyam (private communication) and a thin polycrystalline layer has been found by Brookes (private communication) on molybdenum.



(*a*) ×290. (*b*) ×900.

Fig. 4. Spark-planed aluminium surface.

6. Conclusions

Spark cutting and machining with a rotating slotted cathode is an invaluable method for the production of accurate, strain-free shapes in many metals, either for single-crystal fabrication or for metallographic preparation.

The technique has been used to produce a wide variety of specimens including those for ultrasonic elastic constant measurements, which require accurately oriented strain-free surfaces parallel to within a few microns; single-crystal wafers, for subsequent electrolytic thinning and transmission electron microscopy; single-crystal tensile specimens; single-crystal picture frame specimens for magnetic measurements and many other applications.

* The micrographs reproduced here were obtained with the scanning electron microscope at the Engineering Laboratories, Cambridge University.

References

JAMES, J. A., and MILNER, C. J., 1953, *J. Sci. Instrum.*, **30**, 386.

LAZARENKO, B. R., 1946, British Patent Specification 637,793.

Use of ring-shaped neutron sources in a research reactor

by P. N. COOPER, B.Sc., Ph.D., A.Inst.P., K. FIRTH, B.Sc., Ph.D., A.Inst.P., and K. G. STEPHENS, B.Sc., Ph.D., A.Inst.P., Associated Electrical Industries Ltd., Aldermaston, Berks.

MS. received 14*th October* 1960, *in revised form* 30*th November* 1960

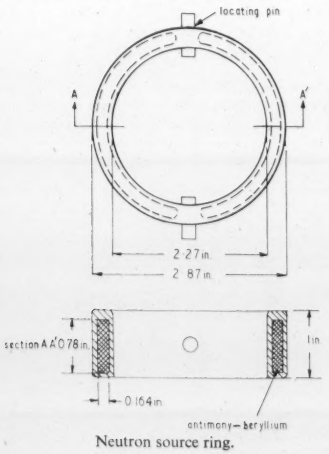
Abstract

Measurements have been made of the effectiveness of the ring shaped neutron sources for the MERLIN *reactor. The emission rate of one neutron source was measured absolutely at the National Physical Laboratory and the remaining sources were compared with this standard. A neutron to gamma ratio of* $(0{\cdot}7 \pm 0{\cdot}1) \times 10^6$ *n/s per curie of* ^{124}Sb *was obtained. The effectiveness of the sources was found by measuring the power generated in a sub-critical core. This result was in reasonable agreement with a calculation, based on two-group diffusion theory.*

1. Introduction

In the MERLIN reactor (Allibone *et al.* 1958) separate neutron sources are fitted to each fuel element to minimize the possibility of the reactor becoming critical without any source neutrons. The ideal position to place the sources for maximum utilization of the neutrons is within the reactor core, but with the MERLIN core this arrangement was not possible.

The sources for the reactor are rings (see Figure) which fit on to bayonet slots on the bottom of each fuel element; thus they do not interrupt the flow of water through the fuel. Each source comprises a pure aluminium shell filled with approximately 36 g of a compressed mixture, 80% by weight of antimony and 20% by weight of beryllium. Sources are fitted to elements under water and, when the fuel elements are in position in the core lattice plate, the mid-plane of each source is 8·2 cm below the active fuel.



Neutron source ring.

2. Measurements of source strength

The total neutron emission of one source, which had been irradiated for eight weeks in a nominal thermal neutron flux of 10^{12} n cm^{-2} s^{-1} in BEPO, has been measured by the National Physical Laboratory. They used the method of activation of a $MnSO_4$ solution, and the neutron emission corrected to the end of the irradiation was $(3{\cdot}5 \pm 0{\cdot}2) \times 10^6$ n/s. The associated gamma-radiation from ^{124}Sb was $4{\cdot}9 \pm 0{\cdot}8$ c, as measured with a calibrated gamma-ray dosimeter. Thus the neutron emission from this type of source is $(0{\cdot}71 \pm 0{\cdot}12) \times 10^6$ n/s per curie of ^{124}Sb.

Relative measurements of the source strengths are performed under 6 ft of water in the reactor tank. The source ring is placed in a jig in a fixed position relative to a BF_3-filled proportional counter. Counting rates from the detector for each source are compared with the counting rate from a 'standard' source, which itself was compared with the source whose emission was measured absolutely. The largest spread in the strengths of sources irradiated together is about $\pm 10\%$ from the mean.

3. Effectiveness of the sources in the reactor

The reactor control instruments have been calibrated in terms of reactor power in watts from relative and absolute neutron flux distributions. These measurements were made with gold and manganese foils at powers up to 20 w. The core, during these measurements, was a 5×4 array of 19 fuel elements, each containing 140 g of ^{235}U, and one fuel element containing 40 g of ^{235}U. With an estimated negative reactivity of 1% and a total neutron source emission of $3{\cdot}3 \times 10^7$ n/s, the core power was found to be $2{\cdot}1 \pm 0{\cdot}3$ mw. This is much lower than initially estimated (Firth *et al.* 1958). A new theoretical estimate, described in § 4, predicted a power of 2·7 mw for the same source emission and negative reactivity.

4. Theoretical estimate of reactor power

A one-dimensional two-group diffusion theory calculation was made in order to find the fraction of source neutrons absorbed in ^{235}U. The system adopted for the calculation consisted of an infinite water reflector, a source plane emitting one neutron per second, 8·2 cm of water, and an absorbing core region 60 cm high plus an extrapolation distance of 7·65 cm to allow for the top reflector. The effect of horizontal leakage from the core was included by means of a buckling correction factor.

The flux distribution of thermalized source neutrons ϕ_{2s} was obtained in order to compute the number of these neutrons captured in the ^{235}U. It was found that this flux distribution was strongly biased towards the neutron source plane. These neutrons are therefore captured in a region of low statistical weight and will be less effective than the slowed down fission neutrons, which have an approximately cosine distribution in a highly multiplying core. Sub-critical measurements have shown that this is true for cores with multiplications greater than 20 (Kerridge, private communication). An attempt was made to estimate the magnitude of this effect by two-group perturbation theory. A separate calculation for a critical core yielded the real thermal flux ϕ_{2c} and adjoint fast flux ϕ^*_{1c}. Hence, the effectiveness β is given by:

$$\beta = \frac{\int_{\text{core}} \phi_{2s}\phi^*_{1c}dz}{\int_{\text{core}} \phi_{2c}\phi^*_{1c}dz} \cdot \frac{\int_{\text{core}} \phi_{2c}dz}{\int_{\text{core}} \phi_{2s}dz}.$$

The effective source strength is therefore:

$$S_{\text{eff}} = \beta\Sigma_2 \int_{\text{core}} \phi_{2s}dz.$$

The two-group constants used in the calculations are given in the Table. Fast diffusion coefficients and ages were obtained from three-group data given by Deutsch (1957). Age was assumed to be proportional to lethargy in the middle energy group, but the diffusion coefficients were taken directly from the middle group, since the scattering cross sections are approximately constant in this energy range.

The effectiveness β was found to be 0·566 and S_{eff} to be 0·032. The multiplication in a core, which is 1% sub-critical, is 100. Assuming, as in the power calibration, that $3{\cdot}3 \times 10^7$ fissions per second are equivalent to 1 mw, the calculated power for a total source strength of $3{\cdot}3 \times 10^7$ n/s is 2·7 mw.

Two-group constants

Diffusion parameter	*Symbol*	*Water*	*Core*
Age (Sb–Be neutrons)	L_1^2	6·65 cm^2	12·62 cm^2
Fast diffusion coefficient	D_1	0·608 cm	0·833 cm
Diffusion area	L_2^2	8·44 cm^2	3·07 cm^2
Thermal diffusion coefficient	D_2	0·161 cm	0·228 cm
Uranium cross section	Σ_{25}	—	0·0576 cm^{-1}
Buckling correction	α^2	0·008 156 cm^{-2}	

5. Conclusions

The ratio of neutron output to gamma-ray output for an antimony–beryllium neutron source in the form of a ring is $(0{\cdot}7 \pm 0{\cdot}1) \times 10^6$ n/s per curie of ^{124}Sb.

The reactor core power due to the neutron sources is in reasonable agreement with that predicted by two-group diffusion theory. In calculating the effectiveness of the sources it is necessary to take into account that the neutrons from the source are captured in the core in a region of low statistical weight. Two-group perturbation theory is adequate for predicting this effect.

Acknowledgments

The authors wish to thank Dr. A. J. Salmon for his advice, Mr. E. J. Axton of the Applied Physics Division, National Physical Laboratory, for measuring the absolute emission of one neutron source, the Staff of the Isotope Division, and the Reactor Division at the Atomic Energy Research Establishment, Harwell, who arranged and carried out the irradiation of the source rings, and Miss M. Cresswell for carrying out the computational work. The authors are also indebted to Dr. T. E. Allibone, Director of the Research Laboratory, Associated Electrical Industries, for permission to publish this paper.

References

ALLIBONE, T. E., CHICK, D. R., FIRTH, K., MILLAR, B., MUNRO, I., and SALMON, A. J., 1958, *2nd Int. Conf. on the Peaceful Uses of Atomic Energy*. Paper A/Conf. 15/P/313 (Geneva: United Nations Publication).

DEUTSCH, R. W., 1957, *Nucleonics*, **15**, (1) 47.

FIRTH, K., KERRIDGE, M., MATHAMS, R. F., SALMON, A. J., STEPHENS, K. G., TROTT, A. J., YOUNG, J. D., 1958, *2nd Int. Conf. on the Peaceful Uses of Atomic Energy*, Paper A/Conf. 15/P/1461 (Geneva: United Nations Publication).

Temperatures under steadily heated floors

by I. G. DONALDSON, Dominion Physical Laboratory, Department of Scientific and Industrial Research, Lower Hutt, New Zealand

MS. received 24*th January* 1961

Abstract

The temperature distribution below the heated floor of a kiln has been investigated both theoretically and by model experiments. The aim is to determine how much insulation is required to keep the temperature of the ground below that point (perhaps as low as 40° C*) at which it tends to dry out and subside. The results show that a considerable thickness of insulation is necessary to achieve this, but that much less insulation is needed if a thin layer of high conductivity is placed below the insulation. Iron reinforcing mesh could act as a high conductivity layer. It also appears that there is an advantage both in cost and in effectiveness in making the insulation thinner toward the sides of the kiln. The conclusions also apply to the insulation of the floor of cold storage rooms and skating rinks where the aim is to prevent 'frost heave'.*

Introduction

THE drying out of clay ground under kilns can lead to subsidence. This usually leads to unsatisfactory operating conditions and a heavy maintenance (or rebuilding) programme. Some knowledge of the temperatures in the ground under these kilns and of the way in which layers of different conductivity below the kiln floor modify these temperatures may lead to a type of kiln construction that minimizes the probability of subsidence.

Under cold storage rooms and ice skating rinks, etc., 'frost heave' may occur if the ground is not sufficiently protected. With appropriate modification, the results and suggestions given in this paper would apply to the floor construction of such buildings.

Ward and Sewell (1950) discuss this problem in some detail, but the theoretical method used by them is restricted by the assumption that the temperature of the kiln floor and the temperature of the ground surface are both constant over the surfaces, and by their assumption that there is no component of the heat flux parallel to the ground surface in either the insulating layers or in any of the surrounding ground that is above the level of the floor. The first assumption means that temperature determinations under the outside walls of the kiln could be considerably in error. Both assumptions restrict the method to situations in which the horizontal heat flow in the floor is not important. The present theoretical study avoids these limitations by treating the insulating layers as being elliptical in section and of finite thickness. Kilns are, however, usually built with a uniform thickness of insulant below the floor, so these predictions will apply only approximately to existing kilns, but in the model experiments the design of existing kilns is imitated more closely.

Approximations in the theoretical treatment

In the present theory it has been necessary to assume that all the floors are of infinite length. Ward and Sewell (1950) give a method of deducing the approximate value of the temperature under the centre of the insulation for other shapes (square, circular and rectangular) from the results for the long-rectangular shape. Hence floors of other shapes will not be discussed in this paper.

Particular examples of heated floors that have come to the writer's notice in industrial problems are illustrated in Fig. 1(*a*), (*b*) and (*c*). These are (*a*) a flat heated floor at ground level; (*b*) a flat heated floor below ground level; and (*c*) a heated floor with a flue for hot gases beneath it. Some insulation is present in each case. In order to apply the mathematical method used here it has been necessary to assume that these designs may be represented by the sections shown in Fig. 2(*a*) and (*b*) without greatly upsetting the

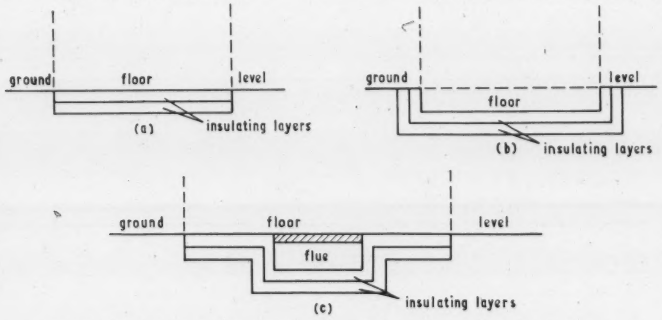


Fig. 1. Typical practical floor arrangements with possible insulation layers shown.

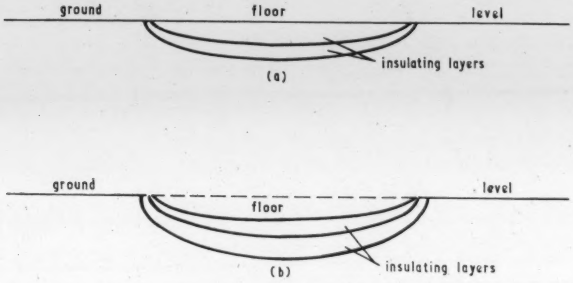


Fig. 2. Floor arrangements as modified for theory.

temperature pattern. Here the floor surface and the surfaces of any insulating layers are assumed to be sections of confocal, elliptic cylinders, the exterior ground surface is taken to be the extension of the diametral plane through the foci. This modification means that the thickness of insulation is greater at the centre than at the edges and also that the insulation extends slightly past the edge of the floor.

When the floor level is below the ground level, as in Fig. 1(*b*), the floor surface in the theoretical model is represented as being curved, as in Fig. 2(*b*), but it is not thought that this difference in shape should affect the predictions by more than a few per cent. When there is a flue beneath the floor as pictured in Fig. 1(*c*), the theoretical model would again be taken to be like Fig. 2(*b*), but it would be possible to represent the flue by a zone of somewhat higher or lower temperature near the middle of the floor. The present theory is able to take account of any prescribed variation of temperature or of downward heat flux over the floor of the kiln. It would also be possible to prescribe the temperature or the heat flux over the ground surrounding the kiln, but it seems more realistic to assume that the heat lost from the ground surface is carried away by convection or radiation. These processes are summarized by an 'edge conductivity' that is considered in the next section.

Edge conductivity

At the ground surface the heat is dissipated by conduction and convection to the air, and by radiation. Heat conduction through the air may be ignored. Thus, heat transfer by convection from a horizontal surface to the air can be described by the equation (Pallot 1954)

$$H = (1 + 0{\cdot}32V)(T_s - T_a) \tag{1}$$

where V is the wind speed (m.p.h.), T_s is the surface temperature and T_a is the air temperature (both in degrees Fahrenheit) and H is the heat loss by convection (Btu ft^{-2} hr^{-1}). The equation is valid for windspeeds in the range from 1 to 10 m.p.h.

As regards radiation losses, a body at an absolute temperature T_1 surrounded by a black body at a temperature T_0 (absolute) will lose heat by radiation at the rate

$$H_2 = \sigma E(T_1^4 - T_0^4) \tag{2}$$

where σ is the Stefan–Boltzmann constant and E is the emissivity of the surface. If $T_1 - T_0$ is small then (2) may be reduced to

$$H_2 = 4\sigma E T_0^3(T_1 - T_0); \tag{3}$$

hence, assuming that neighbouring buildings, trees, etc., are at the same temperature as the air, here taken to be 68° F, the radiative heat loss for most soils and materials of similar texture will be

$$H_2 = 0{\cdot}9(T_s - T_a)\text{Btu ft}^{-2}\,\text{hr}^{-1}. \tag{4}$$

T_s and T_a are now in degrees Fahrenheit.

Thus the total heat loss from the ground surface will be, for ground of conductivity K,

$$H = (1{\cdot}9 + 0{\cdot}32V)(T_s - T_a)\text{Btu ft}^{-2}\,\text{hr}^{-1}. \tag{5}$$

This loss must, of course, be provided by conduction through the ground at a rate $K\partial T/\partial N$, where K is the conductivity and $\partial T/\partial N$ is the temperature gradient normal to the surface, close to the surface. Thus

$$\frac{\partial T}{\partial N} = \frac{1}{K}(1{\cdot}9 + 0{\cdot}32V)(T_s - T_a). \tag{6}$$

The factor of proportionality will be called the 'edge conductivity' h,

$$h = \frac{1}{K}(1{\cdot}9 + 0{\cdot}32V). \tag{7}$$

In what follows it will be convenient to refer all temperatures to that of the air so that a temperature T means the number of degrees by which the temperature exceeds that of the air. The boundary condition (6) at the ground surface is therefore

$$\frac{\partial T}{\partial N} = hT \tag{8}$$

where h is the edge conductivity expressed by Eqn (7). Temperatures quoted in later tables are this temperature T and some 15 deg C should be added to give temperatures referred to the ice point.

It will be noted from Eqn (7) that the edge conductivity is inversely proportional to the bulk conductivity of the material. The theoretical models sketched in Fig. 3 represent all insulating layers below the kiln floor as extending past the edge of the floor and upward to the ground surface. In making calculations one ought to treat edge conductivity as having a different value at these places merely because the bulk conductivities of the materials are not the same as that of the ground. This would be inconvenient, and because the layers are very thin where they reach the ground surface it has been thought sufficient to use a uniform value of h for all the ground surface outside the kiln. The value used is based on the conductivity of the

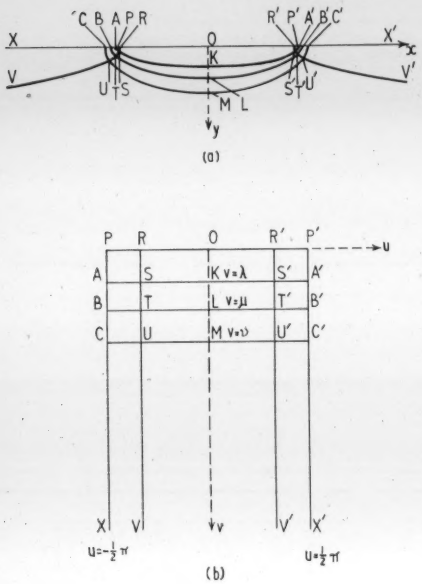


Fig. 3

(*a*). Floor and insulation layers as assumed in the theory.

(*b*). Arrangement illustrated in (*a*) after application of the transformation $z = C \sin \omega$.

ground itself. This is taken to be 1·0 Btu ft^{-1} hr^{-1} deg F^{-1}. Assuming a mean wind speed of 6 m.p.h., h has a value of 4 ft^{-1}.

Theory

The z plane (coordinates $x + iy$) is transformed into an ω plane (coordinates $u + iv$) by the relation $z = C \sin \omega$ where C is half the distance between the foci P and P′ shown in Fig. 3(*a*). The system of confocal elliptic and hyperbolic boundaries in the z plane shown in Fig. 3(*a*) become, in the ω plane, the rectangular boundaries shown in Fig. 3(*b*). Thus any semi-ellipse with foci at P and P′ becomes in the ω plane a finite straight line parallel to the u axis and stretching between $u = -\frac{1}{2}\pi$ and $u = \frac{1}{2}\pi$. Similarly any semi-hyperbola with the same foci and in the same half-plane is transformed to a pair of semi-infinite straight lines ($v \geqslant 0$) parallel to the v axis and sited symmetrically about the v axis. Thus it may be seen that either of the arrangements illustrated in Fig. 2 may be transformed to a semi-infinite rectangle in which the floor surface is now the edge $v = \lambda \geqslant 0$, $|u| \leqslant \frac{1}{2}\pi$; layers of insulation are bounded by parallel lines, $v = \mu$, ν, etc., $|u| \leqslant \frac{1}{2}\pi$; and the ground surface is now represented by the straight lines $u = \pm\frac{1}{2}\pi$.

Laplace's equation is unaffected by this transformation, although some of the boundary conditions are modified. At the ground surface boundary the condition (8) becomes

$$\frac{\partial T}{\partial u} - hT = 0 \quad \text{at} \quad u = -\tfrac{1}{2}\pi \tag{9}$$

and

$$\frac{\partial T}{\partial u} + hT = 0 \quad \text{at} \quad u = \tfrac{1}{2}\pi. \tag{10}$$

At any surface of separation between two layers of different conductivities K and K', if T and T' denote the temperature in the respective layers,

$$T = T' \tag{11}$$

$$K\frac{\partial T}{\partial v} = K'\frac{\partial T'}{\partial v}, \tag{12}$$

and at the surface of the floor the temperature

$T = f(x)$ in the z plane becomes

$$T = F(u) = f\left\{c \cosh\left(\sinh^{-1}\frac{y_1}{c}\right)\sin u\right\} \tag{13}$$

in the ω plane, where y_1 is the depth of the centre of the floor below the line connecting the foci P and P′ (i.e. below ground level).

Let the heated floor have two layers of different conductivities between it and the ground, then in Fig. 3(*b*) the line $v = \lambda$, $|u| \leqslant \frac{1}{2}\pi$ is the floor surface, the lines $v = \mu$ and $v = \nu$, $|u| \leqslant \frac{1}{2}\pi$ are the interfaces between the layers and the lines $u = \pm\frac{1}{2}\pi$, $v \geqslant \nu$ are the ground surface. The conductivity of the upper layer is taken to be K_i that of the lower layer is K_c while the ground has a conductivity K_g. Then in the two layers the temperature is of the form

$$T = \sum_{r=1}^{\infty}\{a_r \exp(\alpha_r v) + b_r \exp(-\alpha_r v)\}\cos \alpha_r u \tag{14}$$

where the α's are the roots of

$$\alpha \tan \tfrac{1}{2}\pi\alpha = h \tag{15}$$

while in the ground, since $T \to 0$ as $v \to \infty$,

$$T = \sum_{r=1}^{\infty} c_r \exp(-\alpha_r v)\cos \alpha_r u. \tag{16}$$

If there are two different insulating layers below the floor there will be two equations of the type (14) and it is convenient to distinguish them by using primed symbols for the lower one.

The five constants, a_r, b_r, a_r', b_r', c_r in these two equations (14) and (16), may be solved using the boundary conditions at the surfaces of separation and at the surface $v = \lambda$, where from Eqns (13) and (14)

$$a_r \exp(\alpha_r\lambda) + b_r \exp(-\alpha_r\lambda) = \frac{2(\alpha_r^2 + h^2)}{\frac{1}{2}\pi(\alpha_r^2 + h^2) + h}\int_0^{\pi/2} F(u)\cos \alpha_r u\,du \tag{17}$$

hence the temperatures in the ω plane are given by

$$T = \sum_{r=1}^{\infty}\chi_r(u)[(K_i + K_c)\{K_c \cosh \alpha_r(\nu - v) + K_g \sinh \alpha_r(\nu - v)\} + (K_i + K_c)\{K_c \cosh \alpha_r(\nu - 2\mu + v) + K_g \sinh \alpha_r(\nu - 2\mu + v)\}] \quad \lambda \leqslant v \leqslant \mu \tag{18}$$

$$T = \sum_{r=1}^{\infty} 2\chi_r(u)K_i\{K_c \cosh \alpha_r(\nu - v) + K_g \sinh \alpha_r(\nu - v)\} \quad \mu \leqslant v \leqslant \nu \tag{19}$$

$$T = \sum_{r=1}^{\infty} 2\chi_r(u)K_iK_c \exp\{-\alpha_r(v - \nu)\} \quad v \geqslant \nu \tag{20}$$

where

$$\chi_r(u) = \frac{2(\alpha_r^2 + h^2)\cos\alpha_r u \int_0^{\pi/2} F(u')\cos\alpha_r u' \alpha u'}{\{\pi\tfrac{1}{2}(\alpha_r^2 + h^2) + h\}[(K_i + K_c)\{K_c \cosh\alpha_r(\nu - \lambda) + K_g \sinh\alpha_r(\nu - \lambda)\} + (K_i - K_c)\{K_c \cosh\alpha_r (\nu + \lambda - 2\mu) + K_g \sinh\alpha_r(\nu + \lambda - 2\mu)\}]}. \quad (21)$$

These results may be used to derive the temperature distribution in the original z plane. The temperature T_i beneath the centre of the upper insulating layer is

$$T_i = \sum_{r=1}^{\infty} 2\chi_r(0) K_i \{K_c \cosh\alpha_r(\nu - \mu) + K_g \sinh\alpha_r(\nu - \mu)\} \quad (22)$$

while T_c the temperature immediately beneath the centre of the second layer is

$$T_c = \sum_{r=1}^{\infty} 2K_i K_c \chi_r(0). \quad (23)$$

If only one layer of insulation is used the results may be obtained from the above by making $\nu = \mu$ (or $K_c = K_i$) and in this case the temperature immediately under the centre of the insulation is

$$T_i = \sum_{r=1}^{\infty} \frac{2(\alpha_r^2 + h^2)\int_0^{\pi/2} F(u')\cos\alpha_r u' \alpha u'}{\{(\alpha_r^2 + h^2)\pi/2 + h\}\{\cosh\alpha_r(\mu - \lambda) + (K_g/K_i)\sinh\alpha_r(\mu - \lambda)\}}. \quad (24)$$

Theoretical results

The finite thickness of walls

The kiln floor was taken to be 21 ft wide with walls 2 ft thick on either side. The overall width was equated to the distance PP′ between the foci in the theoretical model. The central 21 ft was taken to be at 1000° C above the temperature of the outside air. The temperature was supposed to change linearly across the walls, falling from the kiln temperature to the temperature of the outside ground. In some examples (as, for example, that for which the results are illustrated in Fig. 4) the ground temperature outside the kiln was taken to be that of the air; this corresponds to an infinite edge conductivity. In other cases the ground was assumed to have an edge conductivity of 4 ft^{-1} and the outside temperature of the wall was then assessed by a preliminary calculation.

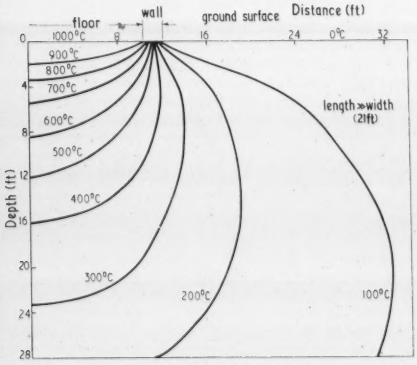


Fig. 4. Temperature pattern under an uninsulated kiln floor. The temperature is assumed to vary linearly through the base of the walls.

The effect of insulation

Calculations were made of the temperature at a point 2 ft below the mid-point of the floor, which was taken to be at the same level as the ground outside the kiln. The ground surface was taken to be at air temperature. Without a special insulating layer the temperature was found to be 900 deg C. With an insulating layer 2 ft thick below the floor and of conductivity 1/15 of that assumed for the ground the temperature at the same point (which would in this case be just below the insulation) was about 190 deg C. This is a considerable reduction but 190 deg C is still not low enough to prevent soil drying and collapsing and the insulating layer could not usually be expected to have a conductivity as low as 1/25 of that of the ground.

To assess what thickness of insulation would be needed to reduce the greatest soil temperature to some specified value, a further series of calculations was made. In this case the ground surface was allowed to have an edge conductivity of 4 ft^{-1}, as suggested previously. The floor was supposed to be at the same level as the outside ground. The results are graphed in Fig. 5, showing the ratio of temperatures at two points below the centre of the kiln, one being on the kiln floor, T_0, and one being in the ground just below the insulation, T_i. The ratio T_i/T_0 is plotted against the ratio of thickness of insulation y to width of kiln $2c$. Curves are plotted for different insulating materials as defined by the ratio of the conductivities of the ground and of the insulant K_g/K_i. The best insulants only just achieve a reduction of ground temperature to one-fifth that of the kiln (say 200 deg if the kiln is at 1000 deg) when the thickness is about one-eighth of the width of the kiln (3·1 ft of insulation for a kiln 25 ft wide, including walls).

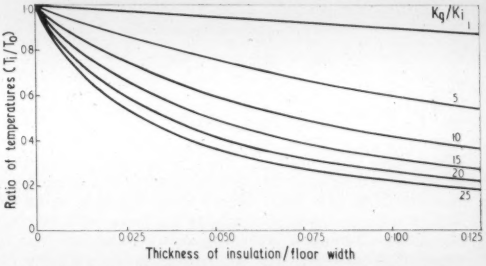


Fig. 5. The variation of temperature beneath the centre of a single layer of insulation with thickness of insulation and the ratio of thermal conductivities of insulation and ground.

Further protection could be obtained by constructing the kiln in such a way that the area beneath the insulation is ventilated, but this has been fully discussed by Ward and Sewell (1950) and is outside the scope of this paper.

The effect of a conducting layer

A consideration of the temperature pattern in Fig. 4 indicates that the heat mainly flows downward, outward, and finally upward, to the ground surfaces on either side, where it is dissipated by radiation and convection. This suggests that a more rapid heat dissipation and a lower temperature

in the ground could be achieved by introducing a thin but highly conducting layer below the insulation. Such a possibility has been studied by the use of Eqn (22) taking λ as zero. In this equation K_i is the conductivity of the insulating layer, and K_c is that of the conducting layer while T_i is the temperature at the mid-point of the interface. The effect of using a thin conducting layer of conductivity eight times that of the ground is illustrated in Fig. 6. The thicknes

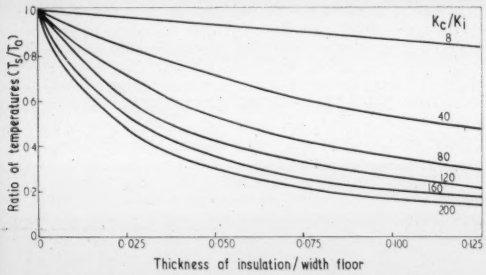


Fig. 6. The variation of temperature under the centre of a combined insulating–conducting layer with thickness of insulating layer. The thickness of the conducting layer is 1/40 width of floor and the conductivity of the layer is 8 times the conductivity of the ground.

of this conducting layer is taken to be 1/40 of the width of the floor.

A convenient way of introducing a conducting layer would be to embed iron reinforcing mesh in shingle below the insulating layer. What matters is the total cross section of iron running across the kiln. The data in Fig. 6 correspond to an iron section of 5·1 in^2 per foot length of kiln, assuming the kiln to be 25 ft wide. If more than one layer of iron rods or mesh were needed to achieve this there would be no theoretical advantage in separating the successive meshes by layers of shingle.

The advantage of using a conducting layer is the resulting economy in insulant. Table 1 illustrates the saving in

Table 1. *The saving of insulation due to the use of a conducting layer. Floor width* 20 *ft, floor temperature* 1000 *deg C above air temperature*

Temp. under insulation	*Cond. insuln. / Cond. ground*	*Thickness of insulation*	*Saving in insulant*
		Insulation alone	
200° C	1/15	3 ft 6 in.	
200° C	1/25	2 ft 2 in.	
100° C	1/15	6 ft 11 in.	
100° C	1/25	4 ft 3 in.	
		With 4 *in. conducting layer* (*cond.* = 12 × *cond. ground*)	
200° C	1/15	2 ft 4 in.	14 in.
200° C	1/25	1 ft 7 in.	7 in.
100° C	1/15	4 ft 6 in.	29 in.
100° C	1/25	3 ft 1 in.	14 in.

different circumstances. Two insulants are considered, one good and one very good. Two amounts of temperature reduction are considered, to one-fifth of the kiln temperature or to one-tenth. The use of a conducting layer saves 7 in. of the very good insulant or 14 in. of the good one if the temperature is being reduced to one-fifth of that of the kiln. The necessary thickness of insulant and the saving that follows the use of a conducting layer are about doubled if one aims to reduce the ground temperature to one-tenth of that of the kiln. It is likely, however, that the effect of a conducting layer may be considerably greater than these theoretical estimates. The reason is that the theoretical model only allows a very thin extension of the conducting layer to reach the ground surface where the heat is dissipated. In any actual construction one would arrange the conducting layer to project some distance beyond the sides of the kiln to allow the heat to dissipate more easily. In this question the model experiments may be a better guide.

It should be pointed out that the conducting layer will not sensibly increase the total heat lost by the kiln. The major loss of heat takes place through the walls and roof.

The theoretical model suggests that when the floor of the kiln is below ground level the necessary thicknesses of insulation or conducting layers will not be very different from the values indicated in Figs 5 and 6. The corrections depend upon the amount that sinh ϕ differs from ϕ, where ϕ is the ratio of the maximum depth of the insulating layer below ground level to the half-width of the kiln. This ratio could typically be 0·3 and the corrections are not then greater than 5%.

Experimental model

It was thought desirable to make model experiments to check the predictions of the theory. In a model one would not be restricted to representing the insulating layers as being elliptical in section; they could conveniently be rectangular slabs in imitation of the usual design. In a model one could also extend any conducting layer beyond the sides of the kiln to improve the dissipation of heat.

The model was made at 1/32 full size so that a metal tank 13½ in. long and 6 in. wide represented a kiln 36 ft long and 16 ft wide. Temperature differences were imitated on a reduced scale too, the tank being filled with water at 50° C, while the ground surface was about 20° C. This difference of 30 deg C could be interpreted as a larger temperature difference, perhaps 1000 deg C, in an actual kiln.

The ground was represented by a bed of coarse dry sand, 15 in. deep, resting on a concrete floor and of overall dimensions 5 ft by 3½ ft. Preliminary tests showed this to be sufficiently extensive. The insulating layer was represented by a rectangular slab of wood (white pine). The conductivities of dry sand and wood are very like the conductivities of dry ground and of insulating concrete. The wood has a rather larger conductivity in the direction of the grain but any errors this would cause were minimized by arranging the grain of the wood to lie horizontally, parallel to the length of the model kiln.

The conducting layer was represented by a mesh of iron wire scaled to represent reinforcing mesh and actually 22 s.w.g. wire at 0·2 in. spacing. It was extended outwards over the ground surface for a distance equivalent to 5 ft beyond the walls of the kiln.

The arrangements of the model are illustrated in Fig. 7. A thin polythene sheet covered the surface of the sand and the embedded wood, the tank of hot water resting on this. Cold water at 20° C was then circulated over the polythene sheet. The purpose was, of course, to keep the sand surface at a known temperature. Since in practice the ground would have a finite 'edge conductivity' and its temperature would, in general, be greater than that of the air, the sand surface

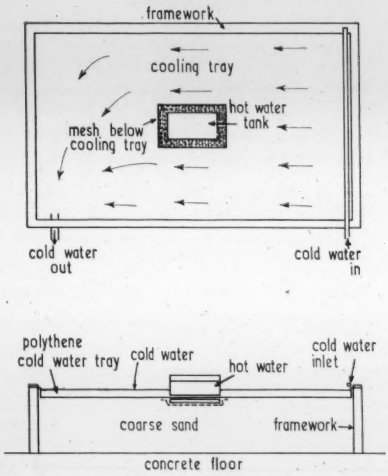


Fig. 7. Plan and elevation of the model (not to scale).

in the model was built up to a level $\frac{3}{16}$ in. above the bottom of the tank; the results would then compare with that for a kiln whose floor was level with the ground outside. This extra layer of sand, since it imitated an edge conductivity and not actual ground, covered all layers such as conducting layers that would actually reach the ground surface outside the kiln.

Temperatures in the sand were measured at points directly below the centre of the 'kiln', by a set of thermocouples whose junctions were placed as accurately as possible at depths of 0, $\frac{3}{8}$, $\frac{3}{4}$ and $1\frac{1}{8}$ in. below the lower surfaces of the wooden insulant. Very fine wires (43 s.w.g.) were used for the thermocouple leads and they were taken out in a direction parallel to the length of the kiln so that they would disturb the temperature pattern as little as possible.

Experimental results

Five different arrangements were used on the model.

1. No insulation other than that given by the polythene sheet.
2. Insulation by wood $\frac{1}{2}$ in. thick.
3. Insulation by wood $\frac{1}{2}$ in. thick on top of the conducting mesh.
4. Insulation by wood 1 in. thick.
5. Insulation by wood 1 in. thick on top of the conducting mesh.

In terms of a full-scale kiln 36 ft by 16 ft, the polythene sheet is equivalent to about 1 in. of insulation. The wooden slabs represent on the full scale an insulant of conductivity 0·56 Btu $ft^{-1} hr^{-1}$ and thickness 8 in. or 16 in. The conducting mesh is equivalent to iron reinforcing rods 0·9 in. diameter laid at 6 in. spacing.

The excess temperature of the sand over that of the cold water, as measured on the thermocouples, was scaled up to present the excess temperature of the ground below a kiln kept at 1000 deg C above air temperature. The actual measurements were made at points equivalent to depths 0, 12 in., 24 in. and 36 in. below the surface of the undisturbed ground under the middle of the kiln, but it seemed reasonable to estimate interpolated values at intervals of 4 in. The results in the five experimental arrangements are listed in Table 2. The figures in parentheses are those predicted by the theoretical method for the temperature at the uppermost point in the undisturbed ground.

Table 2. *Excess temperature (deg C) of ground below kiln at 1000° C. Experimental results obtained with a model for a kiln 36 ft long by 16 ft wide. Figures in parentheses are estimated from the theory*

	Insulation				
Depth below floor (in.)	*None*	*8 in.*	*16 in.*	*8 in. + cond. layer*	*16 in. + cond. layer*
0	910 (910)				
8	770	520 (520)			
12	720	440		400 (390)	
16	680	395	430 (360)	355	
20	645	365	360	325	210 (240)
24	615	345	325	305	185
30	580	320	285	285	165
36	555	295	265	270	155
48	510	260	230	240	135

The agreement between the theoretical and experimental results is, in general, closer than was expected. There is, however, a noticeable difference with the 16 in. of insulation. The model temperature is then considerably higher than that predicted by theory. Since the theory treats the insulant as being thinner at the sides than at the middle, this suggests that a reduction of insulant towards the sides of the kiln may actually allow the heat to be conducted away more easily by the ground. This arrangement would also economize insulating material and is perhaps worth consideration as a point of practical design for future kilns.

Conclusions

The theoretical method laid out in this paper, although only approximating to actual cases, permits the temperature to be computed at any point within, under or around a steadily heated insulated floor. This theory permits conduction along and through any sub-floor layers to be taken into account and may, if required, be extended to include several different layers. The method has been used here for cases where the temperature is stipulated at the floor surface and where the boundary condition (1) applies at the surrounding ground surface. There is no reason why other boundary conditions should not be used at either surface if required. Floors that are below ground level may also be considered.

It is shown that if the temperature under a heated floor is to be lowered to a small fraction (say, a fifth or less) of the floor temperature (as might be required in high temperature kilns set on ground likely to subside on drying out) a considerable thickness of insulation is necessary. To overcome this, without resorting to ventilation, it has been suggested that a thin layer of good conducting material be inserted beneath the insulation and for such arrangements the temperatures that are likely to occur have been computed.

Because this conducting layer is inserted in the ground below a layer of insulation it will have little effect on the heat loss through the kiln floor. If, for example, a 20 ft square kiln 12 ft high is built with walls 2 ft thick and 16 in. of insulation are inserted beneath the floor, a rough theoretical

treatment shows that the heat loss through the floor is increased by less than 15% by the insertion of a conducting layer while the heat loss from the entire kiln is increased by less than 1%.

A series of experiments have shown that for thin layers of insulation beneath a heated floor the theory gives results that compare well with what would actually occur. For thicker layers, apparently on account of the increased sideways heat flow in the ground due to the removal of some of the insulation at the edges, the theory gives temperatures that are lower than are achieved in practice. It may be advantageous in practice, therefore, to reduce the insulation under the edges.

Acknowledgments

This work was done as part of a project for the Pottery and Ceramics Research Association of New Zealand. The author would like to thank Mrs. J. H. Barnard for assistance with some of the computations.

References

CARSLAW, H. S., and JAEGER, J. C., 1947, *Conduction of Heat in Solids* (Oxford: Clarendon Press), p. 99.

PALLOT, A. C., 1954, *J. Inst. Heat Vent. Engrs*, **22**, 1.

WARD, W. H., and SEWELL, E. C., 1950, *Geotechnique, Lond.*, **2**, 64.

New Books

Moderne Messmethoden der Physik: Teil I Mechanik und Akustik. By F. X. EDER. (Berlin: Verlag der Wissenschaften, 1960.) Pp. xxiv + 696. DM 40.

This is the second edition of a book first published in 1951. It is the first of a series which covers a wide field in the methods of experimental physics. This volume is devoted to techniques in Mechanics and Acoustics, another is available on Thermodynamics, whilst Electricity and Magnetism, Optics and Atomic Physics will be dealt with in due course. A number of books have been published in the last few years on experimental methods in physics but most of them have been heavily biased in the direction of the author's personal interests. The present work escapes this fault and the book ranges widely over both of its main subjects. Its chief use will be as a reference book, each small topic occupying about a page in which the fundamental results are quoted together with a short description of methods and some discussion of principles. This basic information is intended to be amplified by using the numerous references. There are about two thousand of these, some of them quite recent.

The whole book has a thoroughness typical of its country of origin, the text is in German and an English edition would be welcome for there is no comparable book already published.

H. J. PAIN

Proceedings of the 1960 Heat Transfer and Fluid Mechanics Institute. Edited by D. M. MASON, W. C. REYNOLDS and W. G. VINCENTI. (London: Oxford University Press, 1960.) Pp. x + 259. Price 70*s*.

The papers are grouped under the headings Vortex motion, Boundary layer flows, Origins of turbulence, Flows not in thermodynamic equilibrium, Separated flows, Two phase systems, and Thermal radiation problems in space technology. Those engaged in the fluid mechanics and heat transfer field will find something to interest them in several of the papers, although they are on the whole rather highly specialized.

An interesting experimental investigation into the mechanism of boundary layer transition is described in a paper by Hama. In common with other groups working in this field, he uses a vibrating ribbon held close to the surface of a flat plate to induce transition to turbulent flow in the boundary layer, and demonstrates that the perturbation wave so formed rolls up into discrete vortices with their axes perpendicular to the direction of flow. These vortices are subsequently deformed into the 'horseshoe' shapes which have been observed in the transition region downstream of trip wires.

The impact of space technology on boiling heat transfer has produced an interesting experiment which is described in a paper by Steinle. He has measured the heat transfer from a heated wire to a liquid under free fall conditions, and finds that the elimination of natural convection produces very considerable changes in the heat transfer process.

Those concerned with the design of space vehicles may like to know of a theoretical study of cooling fins in which the heat transfer mechanism is radiation from the fin surface. Unfortunately the effect of radiation between neighbouring fins is ignored in the paper, and its usefulness is thereby impaired.

The volume is a useful reference work, but is so specialized that most people will not wish to study more than one or two of the papers.

W. B. HALL

Proceedings of the radioactivation analysis symposium held in Vienna in 1959. Presented by the International Atomic Energy Agency Joint Commission on Applied Radioactivity. (London: Butterworths, 1960.) Pp. 141. Price 30*s*.

The sessions of the programme covered radioactivation analysis from various aspects such as geochemistry, biochemistry and medicine, high alloy steels, quartz and germanium, and tantalum ferro-alloys.

Small signal impedance of a solid core inductance

by F. KOLLAR, Dipl.Eng., Ph.D., and R. D. RUSSELL, M.A., Ph.D., Geophysics Laboratory, Department of Physics, University of British Columbia, Vancouver 8, B.C.

MS. received 6th January 1961

Abstract

The design of a highly stabilized control system for a large electromagnet requires the knowledge of its small signal impedance as a function of frequency. The impedance of such an electromagnet was calculated, taking into account the effects of eddy currents but ignoring non-linearities in the magnetizing curve of the iron. It has been possible to use typical values for the relative incremental permeability and conductivity of iron to obtain a rather good estimate of its impedance over four decades above the lowest frequency for which reactive components of the impedance are significant.

There is a change in slope in the curve of the magnitude of the impedance at a frequency corresponding to a skin depth equal to approximately three-quarters of the radius of cross section of the core. Above this frequency the magnitude of impedance rises at 3 *dB per octave and the phase angle approaches* 45°. *At much higher frequencies the existence of stray capacities causes deviations from this predicted behaviour.*

Introduction

During the construction of a mass spectrometer in the Geophysics Laboratory of the University of British Columbia a suitable circuit had to be developed for regulating the magnet current. In order to obtain stability of the order of 1 part in 100 000 the regulator circuits required negative feedback and rather high loop gain. There exist general design methods for servo systems by which it is possible to obtain well damped transient properties of these circuits. The established methods of linear analysis and synthesis can be used in circuits containing well behaved resistors, condensers and amplifiers.

In the case of a large electromagnet and its current regulator the situation is less simple. Because of the solid core of the magnet one expects both the inductive and resistive components of the coil impedance to change with frequency. If the magnet is considered as a transformer with short-circuited secondary, primary impedance changes with frequency according to the penetration depth of the induced eddy currents are predicted. This skin depth is, for example, about 1 mm for the 120 c/s ripple component of the magnetizing current when typical values are taken for the conductivity of the iron ($10^7\ \Omega^{-1}\ m^{-1}$) and for the relative incremental permeability (250). From such considerations it was suspected that the behaviour of the exciting coils could not be described by a transfer function of an easily manageable type.

It is the purpose of this paper to demonstrate a simple method of calculating the amplitude and phase of the impedance of an iron core magnet. Although this procedure includes only eddy current effects, neglecting hysteresis and non-linearities in the permeability, it has been adequate in our case to predict the impedance characteristics of the magnet over four decades above the lowest frequency for which reactive components of the impedance are significant (corner frequency) with an accuracy not worse than 10% for the amplitude and 5° for the phase. Because of the considerable interest in the use of magnets with solid cores in modern experiments in physics the details of this calculation which are presented here should be of general interest.

The large number of publications concerning eddy currents made a thorough survey of the literature difficult, but a few references were found which seemed to give information related to that wanted here. Scott (1930) made calculations about the decrease of inductance with frequency of coils with solid iron cores and gives curves for practical applications. His results do not fit the observed behaviour of our magnet. Rudenberg (1950) discusses magnetic fields in solid ferromagnetic cores and shows their time constants at switching. Frauenberger (1953) deals with different approximations to establish relations between magnetizing current and field in the case of a special arrangement unlike any real magnet. He shows different expressions valid for low and high frequencies. Tuschak (1955) points out errors in Frauenberger's approach.

None of these references suggest that impedance–frequency relations are available in analytical or tabular form to predict the impedance of this magnet and it was decided to find an independent solution. The use of frequency as the independent variable was thought to be suitable because of easy comparison with the directly measurable impedance characteristics. These calculations are now presented here.

Impedance calculations

To determine the impedance of the coil let us assume an arrangement of a toroidal iron core with a circular solid cross section and assume an air gap in the torus (Fig. 1).

The inductance of such a coil can be calculated as $L = N^2/\mathscr{R}$ where N equals the number of turns in the coil and $\mathscr{R}$ equals the reluctance around the magnetic path. This reluctance is the sum of reluctances determined by the iron and the air gap. For frequencies at which the eddy currents in the core become important the reluctance of the iron will vary depending on frequency and will be a complex quantity. In consequence the impedance at the coil terminals will show resulting deviations both in its real and imaginary parts.

Because the air gap reluctance is approximately constant it is only necessary to derive the iron reluctance as a function of frequency. Therefore let us first examine the simple case when there is no air gap in the torus. The symbols used are as follows: E = electric field, H = magnetic field, c = radius of cross section of torus, r = distance from centre of

cross section of torus, σ = conductivity, μ = permeability, ω = angular frequency, I = current, N = number of turns, n = number of turns per unit length of magnetic path, $\mathscr{R}_A$ = reluctance of air gap, $\mathscr{R}_0$ = reluctance of iron core at

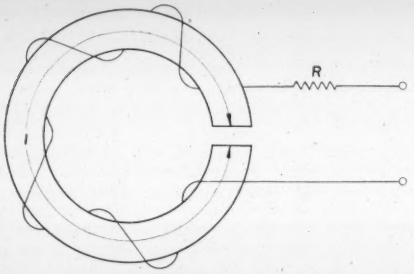


Fig. 1. Assumed model on which the calculations of the impedance were based.

zero frequency calculated with the incremental permeability, R = d.c. resistance of windings.
From Maxwell's equations

$$\nabla \times \mathbf{E} = -\mu\dot{\mathbf{H}} \text{ and } \nabla \times \mathbf{H} = \sigma\mathbf{E}$$

where the displacement currents have been assumed to be negligible in metallic conductors.

Combining these equations there results:

$$\nabla \times \nabla \times \mathbf{H} = -\mu\sigma\dot{\mathbf{H}}$$

$$\nabla(\nabla \cdot \mathbf{H}) - \nabla^2\mathbf{H} = -\mu\sigma\dot{\mathbf{H}}.$$

But within the iron there are no free magnetic poles and $\nabla \cdot \mathbf{H} = 0$ and therefore

$$\nabla^2\mathbf{H} = +\mu\sigma\dot{\mathbf{H}}.$$

Taking $\mathbf{k}$ as a unit vector in the direction of the circular axis of the torus, then

$$\mathbf{H} = H\mathbf{k}$$

and

$$\nabla^2 H = \mu\sigma\dot{H}.$$

If the radius of the torus is sufficiently large compared with c, the assumption of cylindrical symmetry will introduce only minor errors and therefore

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dH}{dr}\right) = \mu\sigma\dot{H}$$

or

$$\frac{d^2H}{dr^2} + \frac{1}{r}\frac{dH}{dr} = \mu\sigma\dot{H}.$$

The variables are separated by writing $H = \mathrm{G}(r)\mathrm{K}(t)$.

Therefore

$$\frac{\mathrm{G}''(r)}{\mathrm{G}(r)} + \frac{1}{r}\frac{\mathrm{G}'(r)}{\mathrm{G}(r)} = \mu\sigma\frac{\mathrm{K}'(t)}{\mathrm{K}(t)} = \lambda \text{ (a constant)},$$

from which $\mathrm{K} = e^{\lambda t/\mu\sigma}$ $\mathrm{G} = \mathrm{J}_0\{r(-\lambda)^{\frac{1}{2}}\}$

and

$$H = ae^{\lambda t/\mu\sigma}\,\mathrm{J}_0\{r(-\lambda)^{\frac{1}{2}}\}$$

where J_0 is the Bessel function of order zero. Applying the boundary condition that $H = nIe^{j\omega t}$ at $r = c$:

$$nIe^{j\omega t} = ae^{\lambda t/\mu\sigma}\,\mathrm{J}_0\{c(-\lambda)^{\frac{1}{2}}\}$$

whence $a = nI/\mathrm{J}_0\{c(-\lambda)^{\frac{1}{2}}\}$ and $\lambda = j\omega\mu\sigma$.

Therefore

$$H = nIe^{j\omega t}\,\mathrm{J}_0\{r(-j\omega\mu\sigma)^{\frac{1}{2}}\}/\mathrm{J}_0\{c(-j\omega\mu\sigma)^{\frac{1}{2}}\}.$$

Now

$$\Phi = \int_0^c \mu H 2\pi r dr$$

$$= \frac{2\pi\mu nIe^{j\omega t}c\,\mathrm{J}_1\{c(-j\omega\mu\sigma)^{\frac{1}{2}}\}}{(-j\omega\mu\sigma)^{\frac{1}{2}}\,\mathrm{J}_0\{c(-j\omega\mu\sigma)^{\frac{1}{2}}\}}. \quad (1)$$

The magnetic reluctance

$$\mathscr{R} = NIe^{j\omega t}/\Phi$$

$$= \frac{I(-j\omega\mu\sigma)^{\frac{1}{2}}\,\mathrm{J}_0\{c(-j\omega\mu\sigma)^{\frac{1}{2}}\}}{2\pi c\mu\,\mathrm{J}_1\{c(-j\omega\mu\sigma)^{\frac{1}{2}}\}}$$

$$= \mathscr{R}_0\frac{\eta(-j)^{\frac{1}{2}}}{2}\frac{\mathrm{J}_0\{\eta(-j)^{\frac{1}{2}}\}}{\mathrm{J}_1\{\eta(-j)^{\frac{1}{2}}\}}$$

where $\mathscr{R}_0 = l/\pi c^2\mu$ and $\eta^2 = \omega\mu\sigma c^2$.

As $\eta \to 0$ the reluctance approaches $\mathscr{R}_0$ and as $\eta \to \infty$ the reluctance approaches $\frac{1}{4}j^{\frac{1}{2}}\eta\mathscr{R}_0$.

Thus it is seen that at high frequencies the reluctance of the iron is increasing with the square root of the frequency and has both a real and imaginary component as the eddy currents lag behind the primary current by 45°.

In the absence of an air gap and where the d.c. resistance of the primary winding is negligible, the complex inductance

$$L = \frac{2N^2\mathrm{J}_1\{\eta(-j)^{\frac{1}{2}}\}}{\mathscr{R}_0\eta(-j)^{\frac{1}{2}}\mathrm{J}_0\{\eta(-j)^{\frac{1}{2}}\}}$$

and the impedance

$$Z(\omega) = j\omega L$$

$$= \frac{2j^{3/2}N^2\eta\mathrm{J}_1\{\eta(-j)^{\frac{1}{2}}\}}{\mu\sigma c^2\mathscr{R}_0\mathrm{J}_0\{\eta(-j)^{\frac{1}{2}}\}}$$

$$= \frac{jN^2\eta^2}{\mu\sigma c^2\mathscr{R}_0\mathrm{f}(\eta)} \quad (2)$$

where

$$\mathrm{f}(\eta) = \frac{\eta(-j)^{\frac{1}{2}}\mathrm{J}_0\{\eta(-j)^{\frac{1}{2}}\}}{2\mathrm{J}_1\{\eta(-j)^{\frac{1}{2}}\}}.$$

Adding the reluctance of the air gap and the finite resistance of the primary winding there results

$$Z(\omega) = \frac{jN^2\eta^2}{\mu\sigma c^2\{\mathscr{R}_A + \mathscr{R}_0\mathrm{f}(\eta)\}} + R \quad (3)$$

which gives the predicted impedance of the magnet at frequencies low enough so that capacities can be neglected. Tables of functions by Jahnke and Emde (1945, p. 266) give the phase and magnitude of the ratio of Bessel functions required here.

Equation (3) assumes a constant air-gap reluctance. In fact the air-gap reluctance must drop at higher frequencies because the restriction of changes in the magnetic flux to the surface of the iron will result in a reduction in the effective cross section area of the air path. It was thought that this would be an effect of secondary importance and this was borne out by the measurements.

Figure 2 shows the magnitude of impedance (arbitrary units), and the phase angles of a toroidal iron core without air gap plotted as a function of η^2, i.e. twice the squared ratio of the radius of cross section to the skin depth in the core ($2c^2/\delta^2$), which is proportional to the frequency. In this idealized case the d.c. resistance of the windings was assumed to be negligible. The impedance increases linearly with frequency at the beginning showing a phase lag of 90 °. At higher frequencies a proportionality with the square root of the frequency and a 45° phase lag is obtained. The change of slope of the impedance curve occurs at a frequency which corresponds to a skin depth of approximately three-quarters of the radius of the core cross section.

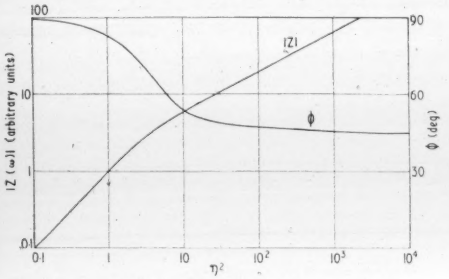


Fig. 2. Form of the impedance curve for a hypothetical iron toroidal coil with no resistance and no air gap.

To extend the calculation for the case when the winding resistance is not negligible and there is an air gap in the core, Eqn (3) can be used. To predict the impedance of the actual mass spectrometer magnet the following numerical values were substituted: $N = 21\,000$ turns, $R = 640\ \Omega$, $\sigma = 10^7\ \Omega^{-1}\text{m}^{-1}$, $c = 0{\cdot}15$ m, $\mu = 250$ (incremental permeability), $\mathscr{R}_A = 2 \times 10^5\ \text{H}^{-1}$, $\mathscr{R}_0 = 0{\cdot}8 \times 10^5\ \text{H}^{-1}$. These two reluctances were calculated from the actual dimensions of the air gap of the magnet and from the dimensions of the iron yoke. The incremental permeability $\mu = 250$ is probably a good approximation for the low carbon content steel used.

Figure 3 shows the magnitude of the impedance and the phase of such a magnet. The corner frequency at which the 640 Ω d.c. resistance changes to the frequency dependent impedance is higher than $\omega = 0{\cdot}2$ radian per second. The skin depth in the core is about $\frac{1}{4}c$ at this frequency which is nine times the corner frequency in Fig. 2. That is the reason why only the slope of one half, corresponding to dependence on square root of frequency, appears in Fig. 2.

Measured impedance of the iron core magnet

There was a preliminary measurement made about one year ago when the magnet was first assembled and the magnet current regulator was designed. The measured impedance versus frequency is shown by Russell and Kollar (1960) in describing certain parts of the mass spectrometer. At that time, because of less adequate measuring facilities, the low frequency results were obtained by recording the magnet current response for a voltage step, and the impedance was numerically evaluated by Laplace transform methods. The results in the frequency domain gave the impedance of the windings quite well for frequencies below 0·2 c/s. Above 2 c/s an audio generator was used for the impedance measurement. Although the magnitude of the impedance was thought to be quite reliable, the fact that the phase angles were decreasing at higher frequencies was puzzling, and to calculate a theoretical impedance-frequency function seemed desirable.

After the predicted impedance plot was obtained, a more precise measurement was desirable to check the theory, which contains several simplifying assumptions not easily justified by calculation. A low frequency signal generator with a frequency range of 0·01 to 1000 c/s was then available but with an output impedance of about 10 kΩ. A transistor power amplifier stage was assembled, using two transistors in the Darlington compound arrangement to obtain the driving signal at a negligibly small output impedance. A resistor of 15 Ω in series with the magnet coil supplied a voltage proportional to the current. The voltage on this resistor and the voltage on the coil were directly connected to

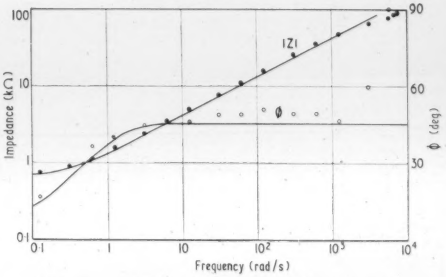


Fig. 3. The impedance curve for a solid iron core inductance with air gap and with finite resistance. The solid curves show the predicted variation calculated from typical electrical parameters for the iron and the measured dimensions and resistance of the coil. The points show experimentally determined values of amplitude and phase of the impedance.

the X and Y input terminals of a Tektronix Type 502 cathode-ray oscilloscope. The dimensions of the ellipses were measured on the screen of the oscilloscope and the impedances and phase angles were determined at 0·01, 0·02, 0·05, 0·1, . . . 1000 c/s. A driving voltage of about 1 volt r.m.s. was used for all measurements. The measured points are plotted in Fig. 3.

The similarity between the measured and the predicted impedances is remarkably good. The phase curves also conform quite well over four decades of the plotted frequency range. Near 1 kc/s a resonance seems to appear due to the stray capacity of the windings.

The impedance curve for a second similar magnet manufactured of Armco iron was also determined. In this case the agreement was less good. The slope of the magnitude–frequency logarithmic plot was 0·57 rather than the predicted value of 0·50 and the phase angle obtained a steady value of 55°. Since the magnets are nearly identical in construction, the difference in behaviour is attributed to the less simple hysteresis curve of the special magnetic iron. However, even in this case the predicted curve is adequate to give an excellent guide for the servomechanism design.

The calculations have assumed a constant value of the permeability and therefore the measurements were made with small signal amplitudes. The actual operating conditions

of the current regulator are similarly determined by the small signal properties of the components. Therefore calculating the impedance in the way shown can supply the necessary information accurately enough, as judged from the measurements, to design the servo loop for optimum stability and speed of response.

Conclusions

A simple calculation method was shown to give the impedance of a magnet with solid iron core as a function of the frequency. The deviations between the predicted and measured frequency characteristics were found to be less than 10% for the magnitude and 5° for the phase. This shows that the neglected hysteresis losses are not important for variations in the current with small amplitudes. Similarly, neglecting the leakage flux of the coils seems permissible. Although the magnet frame has a rectangular cross section, the formulae which were derived for a cylindrical core could be applied.

Therefore it seems to be possible to predict the impedance–frequency characteristics of a magnet using only typical values of magnetic and electric properties of the iron and its physical dimensions, and that the predicted characteristics are sufficiently precise for most practical servomechanism design problems involving the stabilization of the magnetic field.

References

FRAUENBERGER, F., 1953, *Arch. f. Electrotech.*, **16,** 27.

JAHNKE, E., and EMDE, F., 1945, *Tables of Functions*, 4th edn. (New York: Dover Publications).

RUDENBERG, R., 1950, *Transient Performance of Electric Power Systems* (McGraw-Hill).

RUSSELL, R. D., and KOLLAR, F., 1960, *Canad. J. Phys.*, **38,** 616.

SCOTT, K. L., 1930, *Proc. Inst. Radio Engrs, N.Y.*, **18,** 1750.

TUSCHAK, R., 1955, *Arch. f. Electrotech.*, **16,** 392.

New books

Fortschritte der Hochfrequenztechnik. Edited by J. ZENNECK, M. STRUTT and F. VILBIG. (Frankfurt: Akademische Verlagsgesellschaft, 1959.) Pp. xiii + 321. Price DM 42.

This is the fourth volume of a series which reviews progress in the general field of high-frequency techniques. Articles included are: Communication by means of scatter propagation, by J. B. Wiesner (English), Propagation of low and very low frequencies, by H. Poeverlein (German), Modern high-frequency transistors—a survey, by R. L. Pritchard (English), Fabrication techniques for high-frequency transistors, by R. N. Hall (English), Noise in semiconductors, by A. van der Ziel (English), Fluctuation processes in electron beams, by H. W. König and H. Pötzl (German), Delayed conduction in travelling wave tubes and linear accelerators, by K. Pöschl (German), and Properties of getters in electronic tubes, by J. S. Wagener (English). The papers in German have summaries in English and vice versa. The individual articles are well written and authoritative and the standard of production is high. The only doubt about the success of the volume in the reviewer's mind is that the area of coverage in the volume is so wide that the number of potential readers interested in the volume as a whole may be rather small.

C. A. HOGARTH

Computer abstracts. (London: Technical Information Company Ltd.) Published monthly. Subscription rate £25 4*s*. per annum.

Carbon: Proceedings of the Fourth Conference held at the University of Buffalo, New York. (Oxford and New York: Pergamon Press, Ltd., 1960.) Price £7 10*s*.

The *Proceedings of the Fourth Conference on Carbon* contains nearly all the papers presented at the Conference. It is divided into five main sections: Surface properties, Adsorption and reactivity; Electronic properties; Carbonization, graphitization and structure; Mechanical and thermal properties; Carbon technology, friction and wear.

Semiconductor technology. Edited by G. K. T. CONN. (London: Butterworths Scientific Publications, 1960.) Pp. v + 74. Price 10*s*. 6*d*.

This book is a transcript of twelve articles which have been published during 1958–1959 in *Research*. It represents an attempt to present a simple survey of the main topics in semiconductor technology. Among the articles included are some on the preparation of pure materials, on controlled alloying, on transistors and on the problems of handling power. The characteristics of various semiconductor materials are surveyed.

Electrical characteristics of diffused InAs p–n junctions

A study of the electrical characteristics of diffused InAs p–n junctions has been made at 300° K, 196° K and 77° K. At 196° K the current-voltage characteristic is in good agreement with conventional diode theory; deviations from this theory at 300° K and 77° K have been accounted for. There have been reports of photo-voltaic action in InAs junctions in the literature (Tulley and Enright 1954, Hilsum 1957, Lucovsky 1961); however, the rectification properties of these junctions have not previously been discussed.

The junctions used in this study were prepared by diffusing Cd into n-type InAs. The diffusion run was adjusted so that the junction would be 0·001 in. below the surface of the p-region. The n-type InAs was grown in this laboratory; Hall measurements indicated a net donor density of $2{\cdot}4 \times 10^{16}/\text{cm}^3$ at 300° K and $2{\cdot}3 \times 10^{16}/\text{cm}^3$ at 77° K. The average acceptor concentration was estimated from lateral photovoltage measurements (Lucovsky 1961) at 300° K and was found to be of the order of $5 \times 10^{18}/\text{cm}^3$. 1×1 mm samples were cut from the diffused wafers. A broad area indium solder contact was made to the n-region and a smaller indium solder contact of 0·1 mm radius was made to the p-region (see Fig. 1). Previous experience with indium solder contacts has indicated that they are non-rectifying.

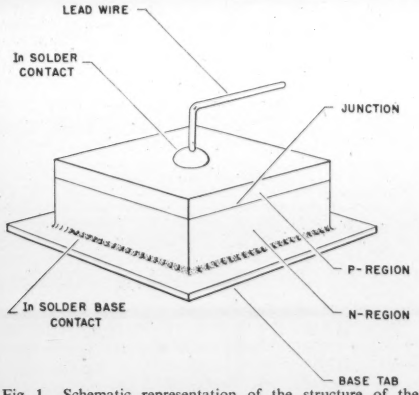


Fig. 1. Schematic representation of the structure of the diffused InAs junctions.

Figure 2 is a plot of the current–voltage characteristic of an InAs junction at 300° K. The rectification ratio at $|V| = 4kT/e$ is found to be 22; conventional diode theory predicts a value of 54. This deviation results from the presence of lateral effects (Lucovsky 1961) which depend on the contact geometry and from pre-breakdown multiplication; the presence of lateral effects reduces the forward current and the pre-breakdown multiplication increases the reverse current. A plot of the logarithm of the forward current against voltage for large forward bias indicates a slope of $e/1{\cdot}30kT$; conventional diode theory based on minority carrier diffusion currents predicts a slope of e/kT (Moll 1958). The deviation from the ideal theory results from non-vanishing lateral resistance which is important due to the contact geometry. The magnitude of the slope of the characteristic is in agreement with an analysis based on this explanation.

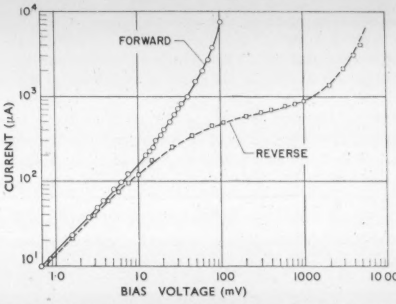


Fig. 2. Current–voltage characteristics at 300° K; InAs junction No. 73.

Further calculations indicate that the generation and recombination of carriers in the transition region can be neglected at this temperature (Sah, Noyce and Shockley 1957).

A plot of the forward characteristic of the same diode at 196° K has a slope of $e/1{\cdot}10kT$. Photoeffect experiments indicate the absence of lateral effects at this temperature; therefore, the contact geometry has no influence on the junction characteristics. If we use the slope of the characteristic at large forward bias as a measure of the extent to which a particular theory is obeyed, we then find that InAs obeys the conventional diode theory at 196° K. Calculations again indicate that the generation–recombination current is small compared to the diffusion current at this temperature.

A plot of the forward current–voltage characteristic at 77° K indicates a slope of $e/2{\cdot}06kT$. Photoeffect experiments indicate no lateral effects at 77° K; therefore, contact geometry is not responsible for the decreased slope. A calculation of the ratio of the diffusion to the generation–recombination current indicates that the latter predominates

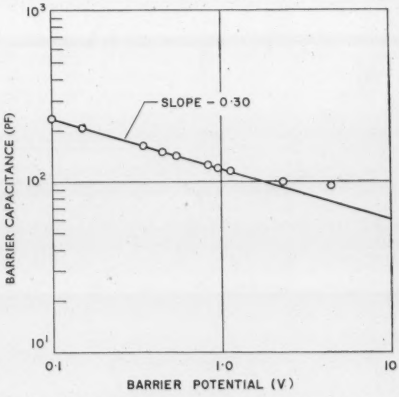


Fig. 3. Reverse barrier capacitance plotted against barrier potential; InAs junction No. 73.

at 77° K and is responsible for the decreased slope of the forward characteristic.

Figure 3 contains a plot of the barrier capacitance as a function of the applied potential for the reverse direction. The voltage scale has been adjusted for the intrinsic barrier potential. The slope of the curve is found to be $-0{\cdot}30$. This value is typical of those obtained in this laboratory for diffused InAs junctions. The failure of the last two experimental points to fall on the curve is attributed to the fact that these points fall in the breakdown region of the characteristic.

The author wishes to thank H. Altemose for his invaluable assistance in the preparation of the experimental samples.

Research Division, G. LUCOVSKY
Philco Corporation,
Philadelphia,
Pennsylvania, U.S.A.

References

HILSUM, C., 1957, *Proc. Phys. Soc.*, **B70**, 1011.

LUCOVSKY, G., 1960, *J. Appl. Phys*, **31**, 1088.

MOLL, J. L., 1958, *Proc. Inst. Radio Engrs*, **46**, 1076.

SAH, C. T., NOYCE, R. N., and SHOCKLEY, W., 1957, *Proc. Inst. Radio Engrs*, **45**, 1228.

TALLEY, R. M., and ENRIGHT, D. P., 1954, *Phys. Rev.*, **95**, 1092.

Notes and comments

International Symposium on Microchemical Techniques

An International Symposium on Microchemical Techniques is being held from 13th to 18th August 1961 at the Pennsylvania State University. The purpose of this symposium is to provide for the interchange of information and ideas among technologists from all parts of the world concerning new methods and techniques or unique applications of microchemical or microanalytical interest.

Further information may be obtained from The International Symposium on Microchemical Techniques Conference Center, The Pennsylvania State University, University Park, Pennsylvania, U.S.A.

2nd International Congress on Cybernetics

The International Association for Cybernetics has just issued the *Proceedings of the Second International Congress on Cybernetics*, held at Namur, Belgium, in September 1958.

This important scientific publication, which includes 80 of the papers read at the Congress, is available from the Secretariat of the 'Association Internationale de Cybernétique, A.S.B.L.,' rue Basse-Marcelle 13, Namur, Belgium. Price 800 Belgian francs for members of the Association and 1200 Belgian francs for non-members.

Journal of Scientific Instruments

Contents of the June issue

PAPERS

Micro-tensile testing machine. By D. M. Marsh.
Basis of verification of metal and glass Sikes hydrometers. By P. H. Bigg.
A time-resolved spectrograph for use with triggered light sources. By F. L. Curzon and J. R. Greig.
Pulse-sampling voltmeter and its application to transistor testing. By L. J. Herbst and J. R. W. Smith.
An instrument to sharpen microtome knives. By A. L. Sims.
Mechanical trace reader. By E. W. Northover.
Tensile testing machine for whiskers. By H. B. M. Wolters and F. W. Schapink.
New quenching circuit of high quenching power and high sensitivity. By B. Kalab.

LABORATORY AND WORKSHOP NOTES

An elastically hinged lever. By E. Eisner.
Two-transistor oscillator for displacement measurements. By L. M. Trémouroux.
Device for protecting recording pens from interference from surgical diathermy. By W. W. Mapleson.
An effective splash trap for a manometer. By F. S. Maguire and A. G. Thomas.
Electroforming as a method of obtaining O-ring grooves. By S. R. Rajagopalan and R. Radhakrishnan.
Simple means of computing two-dimensional steady-state heat flow in composite systems. By D. W. Stops.
A simple trichlorethylene degreasing plant. By P. E. Secker.
Pneumatically driven shutter for high-speed cameras. By J. C. Muirhead and F. L. McCallum.
Simple method of drawing glass fibres. By I. Chavet and B. Davidof.

NOTES AND NEWS

New instruments, materials and tools Notes and comments

The Craftsmanship and Draughtsmanship Competition of The Institute of Physics and The Physical Society.

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.
EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.
REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.
ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611–2.)
SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12s. 6d. each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.
CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

UNWIN BROTHERS LIMITED, WOKING AND LONDON

London symposium on electrical contacts

by M. R. HOPKINS, M.Sc., Ph.D., F.Inst.P., University of Wales, University College of Swansea

Report on the Symposium on Electrical Contacts held by The Institute of Physics and The Physical Society in collaboration with The Institution of Electrical Engineers on 5th–7th April 1961

MS. received 27th April 1961

Abstract

A Symposium on Electrical Contacts was held by The Institute of Physics and The Physical Society, in collaboration with The Institution of Electrical Engineers on the 5th, 6th and 7th April 1961, at the Brunel College of Technology. The scope of the discussion was such as to be of interest to physicists, electrical engineers, metallurgists and chemists from industry and industrial research establishments, from research associations and from universities.

The opening session consisted of an address by Professor F. Llewellyn Jones on 'The physics of electrical contact phenomena', in which he surveyed the various aspects of contact operation, the important problems still outstanding and the methods by which they are now being attacked. The sessions which followed were devoted to 'Principles', 'Fundamental investigations and techniques', 'Contact surfaces', 'Materials and design', 'Non-metallic contacts' and 'Miscellaneous subjects'. At each of these sessions, four or five short papers were presented and followed by discussion.

A Symposium on Electrical Contacts was held by The Institute of Physics and The Physical Society, in collaboration with The Institution of Electrical Engineers on the 5th, 6th and 7th April 1961, at the Brunel College of Technology. The symposium was attended by over two hundred and sixty members, including several overseas visitors. The scope of the discussion was such as to be of interest to physicists, electrical engineers, metallurgists and chemists from industry and industrial research establishments, from research associations and from universities.

The opening session consisted of an address by Professor F. Llewellyn Jones on 'The physics of electrical contact phenomena', in which he surveyed the various aspects of contact operation, the important problems still outstanding and the methods by which they are now being attacked (published in this issue, following paper). The sessions which followed were devoted to 'Principles', 'Fundamental investigations and techniques', 'Contact surfaces', 'Materials and design', 'Non-metallic contacts' and 'Miscellaneous subjects'. At each of these sessions, four or five short papers were presented and followed by discussion. A list of the twenty-five papers presented at the symposium is appended, and the papers are numbered in order of presentation for purposes of reference in this report.

Opening lecture and session on general principles

Problems relating to the design, manufacture and operation of efficient and reliable electrical contacts are of long standing in the electrical engineering and communication industries. Until recent years, research has on the whole been directed mainly to the solution of specific problems and a great deal has been published on empirical studies of special applications. It has become recognized that progress must depend upon a more fundamental approach involving an investigation of the basic physical phenomena associated with the operation of a contact (1). The function of an electrical contact is to close a circuit, to allow a current to pass for a specified time and then to open the circuit. These operations may have to be repeated at certain intervals. Time factors may vary from microseconds to years, potential differences from a small fraction of a volt to thousands of volts, currents from a microampere to thousands of amperes. It is therefore convenient to classify contacts as belonging to one of three main classes; heavy duty contacts, light duty contacts and electrostatic contacts. The first class contains the circuit-breakers used in power installations. This type of contact, in which the problem is mainly that of arc suppression on a large scale, was not considered in the symposium. The contacts used in light engineering and in communication engineering fall into the second class, potential differences being of the order of volts and currents of the order of amperes. In the third class are the contacts which are used in electronic circuits and sensitive electrostatic measuring equipment in which the currents are very small. In light duty contacts failure is due to the disfiguration of the electrodes caused by the transfer of metal from one electrode to the other. The problems that arise in electrostatic contacts are those due to the formation of oxide films and tarnish layers on the electrode surfaces.

It is interesting to consider the three phases of operation of a light duty contact. As the electrodes approach each other a stage is reached at which the electric field between them is high enough to extract electrons from the cathode and to establish conditions under which an arc is formed. This is the well known 'arc at make'. The second phase is that in which contact is made and current flows. Current continues to flow as the electrodes begin to separate until the connection between them is a microscopic bridge of molten metal. When this breaks, we have the third phase, that of a micro-arc between the two electrodes. This is the 'arc at break'.

Much work has been done on the formation, stability and rupture of molten metal bridges between contact electrodes. The shapes of bridges which are stable under the action of surface tension and electrodynamic forces have been calculated and observed. In the opening address, a ciné film

showing sequences of such shapes was shown (1). Although temperature asymmetry in the bridge itself, due for example to the Thomson effect, provides a possible cause of metal transfer, if it is assumed that the bridge breaks at its hottest cross section at a temperature which often approaches that of the boiling point of the metal, the bridge is of central importance in that it sets the stage for the micro-arc which occurs when it breaks. The sudden breaking of the bridge gives rise to a small region containing metal vapour at high pressure, and movement of matter in the form of ions can take place in the high electric field between the electrodes. The bridge between contact electrodes is an example of a conducting system in which the heat lost from the surface can be neglected in comparison with the heat conducted through the system. In this case there is a simple relation between the generalized potential ψ, defined, to include thermoelectric effects, by

$$\mathbf{j} = -\kappa \operatorname{grad} \psi$$

where $\mathbf{j}$ is the vector current density and κ the electrical conductivity, and the temperature θ. This relation, which is a steady state relation, is independent of the geometrical configuration of the system. It is known as the (ψ, θ) relation, and is given by

$$\psi^2 = F^2 - 2\mu$$

where

$$F(\theta, \theta_m) = +\left(2\int_{\theta}^{\theta_m} \frac{\lambda}{\kappa} d\theta\right)^{1/2}, \quad \mu = \int_0^{\psi} P d\psi, \quad P = \int_{\theta_m}^{\theta} \sigma d\theta$$

θ_m is the maximum temperature, λ the thermal conductivity and σ the Thomson coefficient.

If the temperature is uniform, say θ_0, at distances remote from the hottest point, then the potential difference V between the ends of the system is given by

$$\tfrac{1}{2}V = F(\theta_0, \theta_m) = \left(2\int_{\theta_0}^{\theta_m} \frac{\lambda}{\kappa} d\theta\right)^{1/2}.$$

This is a unique relation between V and θ_m which has formed the basis of a method of determining the ratio λ/κ of the thermal to the electrical conductivity of a metal as a function of temperature by observing an artificial bridge consisting of a short fine wire joining two large masses of the same metal (2). The method consists of measuring the temperature of the hottest part of the wire by matching with the strip filament of a standard pyrometer lamp and measuring the potential across the conductor at the same time, over a range of values of temperature and potential, θ_0, the temperature at distances remote from the hottest point, is effectively room temperature. The properties of metals at high temperatures can be investigated in this way and in the case of some metals such as platinum, the wire remains in place even when molten, and measurements can be made well into the molten range. By measuring the current at the same time it is possible to determine λ and κ separately, but this is more difficult.

Some interesting results on thermal instability in electrical contacts (5) depend upon a consideration of the actual variation of electrical and thermal conductivities with temperature in certain cases. Considering two electrodes in contact over a small area, for many conductors, steady conditions can exist only when the current is below a critical value. Greater currents will cause the temperature to rise until some other phenomenon such as melting limits the process. The critical value of the current is associated with an upper limit to the steady temperature which may be created in the body; this may be several hundred degrees below the melting point of the conductor. These conclusions apply to the small regions of contact near the interface of the electrodes of a switch and have interesting implications in the design of electric switches and of fuses.

Since oxide films play a predominant role in the behaviour of contacts, a discussion of the mechanisms of formation of oxide films on metals (3) was appropriate at the symposium. Thermochemical considerations determine whether a metal will form an oxide under particular conditions. Examples were given of the reactions of precious metals, often used as contact materials, with oxygen at a pressure of one atmosphere. Gold is the only metal that does not form an oxide, but this property is shared with certain dilute alloys of other metals in gold. Ag_2O dissociates at about 200° C and Ag_2S, which may be present on silver at room temperature, dissociates at a somewhat higher temperature. PtO_2 volatilizes at an appreciable rate above about 1000° C. Kinetic considerations are of great importance. With thin oxide layers, formed at low temperatures, a strong electric field is responsible for pulling ions across the layer to form new oxide. Various growth–time relationships can be deduced, depending on which step in the process is assumed to be rate-determining. None of these mechanisms is as yet well established. For thick layers (>5000 Å) and higher temperatures, the mechanism of oxidation leading to a parabolic growth–time relationship is well understood and consists of diffusion of ions along a concentration gradient across a protective layer. The species which diffuses is determined by the defect structure of the oxide, and it is possible to reduce oxidation by decreasing the concentration of vacancies or interstitial cations by additions of suitable alloying elements. The oxidation of tungsten is an important example of a more complicated process in which a thin protective layer of a lower oxide grows to a limited thickness next to the metal and then transforms at a constant rate to a powdery outer layer of trioxide, resulting in an initially parabolic growth–time relationship becoming linear at a later stage.

The importance of surface phenomena in non-metals was brought out in a paper on the friction and wear of carbon (4). The conclusion drawn from experimental results was that the low friction of graphitic carbon (and non-graphitic carbon) is due to low adhesion between crystallites and not to low shear strength as is generally supposed. The low adhesion is promoted by adsorption of gases which saturate the forces on the surface of the crystallites, especially the edge forces, and the high friction and wear of carbon brushes which occurs at high altitude arises as a natural consequence of the desorption which then occurs.

Fundamental investigations and techniques

An understanding of metal transfer requires detailed knowledge of the amount of metal transferred from one electrode of a contact to the other. Although this has long received much attention, measurement has always been a matter of some difficulty because of the very small amount of metal transferred in a single operation of a light duty contact. Weighing the electrodes to determine the transfer requires something like 10^6 operations of the contact before an estimation can be made. The gravimetric method also has the disadvantage of determining only the net transfer and not revealing the details of the exchange of metal between the two electrodes. Optical estimation at high magnification also suffers from these disadvantages, which have been over-

come by measuring transfer by using radioactive isotopes (7). If one electrode of a contact has been made radioactive by neutron bombardment in a nuclear reactor, transfer of metal from it to the other electrode can be measured accurately by the use of standard counting techniques. Transfer in the opposite direction is determined by reversing the polarity. In this way the relation between movement of matter in both directions and the circuit inductance has been investigated down to very low values of the inductance—of the order of 10^{-8} H. The relation between the total amount of metal moved, the net transfer and the volume of the molten metal bridge leads to a better understanding of the relative roles of the molten metal bridge and the micro-arc which occurs at the breaking of the contact. In the case of the various metals investigated it was found that there was no range of inductance over which the transfer was independent of inductance.

The result of metal transfer in a light duty contact is the accumulation of excess metal known as a 'pip' on one electrode and the formation of a corresponding crater in the other. Failure of the contact can be due to the mechanical locking of the pip in the crater. Quite apart from any physical theory of transfer, therefore, given that transfer occurs at each operation, the geometrical build up of the pip and crater is of interest. An analysis of this has been made (8) on the basis of the erosion pattern observed after a single operation of the contact, the observable parameters being bridge volume, the fraction of this volume deposited on one electrode and the mean distance to which the bridge material is scattered. In this way relationships are derived which give the rate of growth of the eroded area and the shape of the pip and crater profiles. These are in good agreement with experimental evidence and the theory has a number of practical implications. Metals of high electrical conductivity give rise to sharper pips; also, the larger the radius of curvature of the electrode surfaces the flatter the pip tends to become.

Metal transfer, which normally has a deleterious effect in electrical contacts can be put to good practical use in the process sometimes called 'spark-hardening' (6). In this process a wear resistant skin is generated upon a metal surface by transfer from a vibrating electrode which is smaller than the work-piece and so attains a greater temperature. Experiments were described showing that this temperature difference promoted transfer from electrode to work-piece.

The mechanical deformation of the electrodes of a contact at closure is of considerable importance in relation to the electrical behaviour of the contact. Progress has been made in the study of closure problems by the use of models and by experiments with real contacts (9). Plastic compression and indentation have been investigated by means of models made of two-colour laminated 'Plasticene', the models being of simple geometrical shapes—hemispheres, cones of different apex angles, rectangular blocks. Indentation of a body by a smaller one usually occurs when the smaller is the harder, but it can also occur when the bodies are of identical hardness; its extent is controlled by friction and diminishes with lubrication. The form of the indentation approximates to that of the segment of a sphere, and is almost independent of the original shape of the indenter. Different forces produce indentations which are geometrically similar and so the fracture of an intervening film cannot necessarily be ensured by simply increasing a normally applied force. The flow pattern depends upon the rate of deformation, and for very slow rates the flow does not proceed smoothly. The pattern also depends very strongly on frictional constraints.

Contact surfaces

The session of the symposium devoted to contact surfaces opened with a paper on the cleanliness of precious metal surfaces (10), describing a contact resistance technique for examining the incidence and behaviour of thin adsorbed films of thickness 5 Å to 20 Å on precious metal surfaces and giving the results of contact resistance measurements carried out over the load range 10^{-3} to 10^{2} g. At loadings greater than 1 g it is the bulk properties of the electrode material which determine the behaviour of the contact, whereas below 1 g the surface properties predominate. The results of measurements of metals of the platinum group give a specific film resistance corresponding to film thicknesses of about 5 Å. Film thickness depends on the method of surface preparation. Cold formed surfaces and surfaces of broken sintered components are stable in air for periods of at least a year, in contrast to abraded surfaces which show a progressive deterioration with time. The thickness of films 20 Å thick at 50 mg varies with loading, while the thinner films remain sensibly constant over the range 0·5 to 5000 mg. Low contact resistance, high contact stability and high weld strength are found to go together.

Since the surfaces of contact electrodes are in general covered by oxide and other films and may be contaminated by dust, etc., it is important to examine the effect of such films on electron emission from the cathode, under the action of the electric field in the contact gap as the electrodes approach at the making of the contact (11). This electron emission is the decisive factor in the initiation of the arc at make. The probability of emission of an electron from such a surface as a function of the applied electric field has been measured under a variety of conditions in order to investigate the mechanism of field emission at room temperatures. The technique involves the study of the statistical distribution of the time lag between the application of the field and the emission of an electron. The results show that substantial electron emission occurs at fields of relatively moderate strength (10^{4} to 10^{5} v/cm) compared with the 10^{7} v/cm required by the Fowler–Nordheim theory of emission from clean surfaces. The rate of emission depends upon the state of the cathode surface, and is considerably enhanced by oxide films and dust. The fact that emission takes place at 10^{4} to 10^{5} v/cm rather than 10^{7} v/cm leaves sufficient space between the approaching electrodes for an arc to be established.

The presence of dust can, of course, interfere mechanically with the closure of a contact. This may be discussed (12) in terms of the probability of establishing electrical contact in any particular closing operation. It can be shown that this probability is related in a simple way to the number and size of the contaminating particles, to the mechanical force on the contact and to the nature of the contacting surfaces.

The use of sliding contacts introduces problems of mechanical wear and friction in addition to those of electrical erosion. The behaviour of a typical sliding contact system in which a metal electrode slides over plastic insulating surfaces and metal surfaces in turn has been studied (13). A metallic surface in sliding contact with a plastic surface can pick up plastic material which may be transferred to a second metal surface as the first subsequently slides over it. The plastic acts as a lubricant and the wear of the metallic

surfaces can be substantially reduced without significant interference with electrical contact. With a suitably refined technique, determinations of film thickness can be made by direct weighing and preliminary work has shown that, with polytetrafluoroethylene (PTFE) and nylon, it is possible for palladium, copper and nickel–silver to acquire films having an average thickness of about 250 Å. The process of pick-up does not seem to be wholly mechanical and may be related to polymer formation. It depends not only upon sliding speed and contact force but on the particular plastics and metals involved and on the condition of the metal surfaces. With some combinations of metals and plastics the metal exhibits negative pick-up.

Electrodeposition is an important method of preparation of contact electrode surfaces and there is a good deal of knowledge available on plating conditions, etc. The electrodeposition of precious metals (gold, silver palladium, rhodium and platinum) was discussed (14) in terms of electrolyte formulations, operating conditions, basis metal pre-treatments and properties of deposits, with special reference to the requirements of light duty electrical contacts and the plating of copper-clad laminates.

Materials and design

The session devoted to materials and design was preceded by a review paper on practical aspects of light duty electrical contacts (15), in which the factors which must be taken into consideration in the design and manufacture of such contacts were discussed. These factors are material, electrode shape, surface finish and size and method of attachment. Operational features which must be taken into account are contact loading force, separating force and freedom from bounce. An important factor too is the cost of materials and manufacture. Palladium and certain alloys of ruthenium and platinum and palladium and silver are good from the point of view of freedom from welding, and oxidation of these materials is not so serious as to interfere with the operation. Silver is a good contact material but it tends to weld. A high separating force and freedom from bounce are required.

The practical aspects of applying precious metal finishes to contact surfaces of connector springs, blades and printed circuits were discussed in a paper on finishes for electrical contacts (16). The influences of design, materials and plating processes were discussed and ways and means of selective plating proposed with a view to economy and to extending the use of precious metal contact surfaces. Gold, palladium and composite rhodium coatings on printed circuits have been assessed in relation to the performance of various plating baths and deposit characteristics such as internal stress, hardness and surface resistance. A new process for the deposition of palladium and similar metals was described, and the operation of a typical precious metal plating shop, inspection and testing procedures in the production of reliable electrical contacts were discussed.

An investigation into combinations of silver, hard gold (of various types and hardness values), fine gold and palladium as electrodeposited contact materials was described in a paper on the evaluation of electroplated contacts (17), the field of application being signal transmission circuits. The discussion included consideration of contact resistance under static and sliding conditions, wear resistance, the effect of atmospheric contamination, surface finish coefficient of friction and relative costs.

The general requirements which have to be satisfied in aircraft electrical systems and their implications in switchgear design were considered in a contribution on electrical contacts in aircraft switchgear (18). Experimental work carried out to provide information at the design stage to ensure freedom from high contact resistance and welding during the lifetime of the switch was described.

The difficulties of ensuring satisfactory contact operation at ambient temperatures up to 500° C are related to oxidation, welding and thermoelectric effects (19). Oxidation characteristics of the precious metals at temperatures up to 900° C have been investigated by light load contact resistance measurements. From this point of view gold and platinum are film free and ruthenium is unusual in having a highly conducting oxide. High melting point materials such as iridium and ruthenium have an advantage when welding due to high currents at elevated temperatures is a difficulty. Thermoelectric effects between dissimilar metals, between metals and their oxides and between like metals, at different temperatures, separated by layers of adsorbed vapour were discussed.

The principles of contact design and a comprehensive account of the various factors which must be considered were given in a paper on the selection of electrical contact materials for light duty applications (20). The results of tests on various contact materials including platinum and platinum-group alloys, gold and gold alloys, silver and silver alloys show that life and reliability depend upon surface contact resistance, arc erosion, material transfer and welding characteristics.

Non-metallic contacts and miscellaneous subjects

Granular carbon aggregates are of considerable practical importance, and the heat treatment of such aggregates in various gas atmospheres has been studied (21). In inert gases, in hydrogen and *in vacuo* the electrical resistance either remains constant or decreases as the heat treatment temperature is raised. In ambient gases containing either free or combined oxygen, however, a sharp maximum is obtained in the resistance at a heat treatment temperature of about 530° C. This is attributed to the formation of a non-conducting surface oxide layer which decomposes at higher temperatures.

The behaviour of granular aggregates of carbon and silver carbide were further discussed in a paper on some electrical properties of non-rectifying semiconducting contacts. A single contact, or an aggregate of contacts, in a granular material, display a capacitance and energy loss that vary with frequency. The energy loss can have both a maximum and a minimum value, the location of which depends not only on the intrinsic properties of the material but on granule dimensions and on the structure of the intergranular boundaries. There is also a dependence on the magnitude of the applied field and, in particular, on the magnitude of a superposed steady field. Such a field can produce heating at the boundaries. Theory predicts and experiments confirm that this can cause the small-signal capacitance to assume negative values. Dielectric measurements over a wide frequency range have been made on aggregates of commercial carborundum and heat-treated anthracite. The results suggest the possible existence of a second constituent in each material, and have implications in relation to the ageing of microphonic contacts.

Anyone who has made crude electrical connections by twisting wires together would be interested in the demonstration that this process can be refined and mechanized to give very reliable wrapped or bound solderless connections (23). Hand and power-driven tools have been developed

to do this, and series of tests have shown the suitability of connections of this type in telecommunications equipment.

A paper on electrical contact research in Japan (24) contained a description of results on arc-suppression and on the reduction of contact resistance by electro-chemical treatment of contact electrode surfaces. Research on contact materials includes investigations of Ag–Ni and Ag–metallic oxides produced by the methods of powder metallurgy. An account was also given of the effect of environmental conditions on contact erosion, and of apparatus for electrical measurement in contact research.

The importance of alloy structure in metal transfer in contacts (25) was first realized in experiments with palladium–copper. The change of the structure of an alloy containing 40% copper from the disordered to the ordered state resulted in a reduction in transfer due to the increase in thermal conductivity brought about by the change. Similar effects have been obtained by precipitation hardening of gold–cobalt and gold–nickel alloys. The influence of thermal conductivity on the various possible modes of transfer was discussed.

In the summing up of the symposium (26) reference was made to the fact that this symposium was the second meeting of its kind to be held in this country on the subject of electrical contacts. The first, organized by The Institution of Electrical Engineers was held at Loughborough in 1952. Progress has been made, in the intervening nine years, in fundamental research, in the solution of special problems and in the improvement of manufacturing processes. Ideas on the physical phenomena associated with the operation of light duty contacts which were tentative then have become more fully understood now. More comprehensive and more reliable experimental results are now available as a result of improved techniques and also as a result of a firmer theoretical foundation which provides the basis for further investigation. A great deal more is now known about the properties and behaviour of the metals that are of interest in contacts. Some of the difficulties of the subject are due to the necessity of investigating the interaction between various physical processes which cannot be considered in isolation in the operation of an electrical contact.

List of papers read at the symposium

1. The physics of electrical contact phenomena.
   F. Llewellyn Jones, University College of Swansea (this issue, p. 318).
2. Contact physics and the properties of metals at high temperatures.
   M. R. Hopkins, University College of Swansea.
3. Mechanisms of formation of oxide films on metals.
   O. Kubaschewski and B. E. Hopkins, National Physical Laboratory.
4. An investigation of the friction and wear of carbon.
   J. W. Midgley, A. Strong and D. G. Teer, The English Electric Co. Ltd.
5. Thermal instability in electric contacts.
   F. P. Bowden and J. B. P. Williamson, The Cavendish Laboratory and Tube Investments Ltd. (see *Proc. Roy. Soc.*, **26**, 1).
6. The influence of electrode temperature on transfer of matter in spark-hardening.
   P. E. Watts, Associated Electrical Industries Ltd.
7. Investigation of metal transfer using radioactive isotopes.
   C. H. Jones, University College of Swansea (now at the Institute of Cancer Research, Royal Marsden Hospital).
8. Electrical contact erosion: A theoretical study of pip and crater shape.
   J. Pullen, Electrical Research Association.
9. Closure of a metallic contact.
   A. Fairweather, D. G. M. Shirley and R. E. Fudge, Post Office Research Station, Dollis Hill.
10. The cleanliness of precious metal surfaces.
    H. C. Angus, The International Nickel Co. (Mond) Ltd.
11. Electrical properties of electrode surfaces and surface films in relation to electrical contacts.
    E. B. Pattinson, University College of Swansea.
12. Influence of dust and surface films on the behaviour of electric contacts.
    J. B. P. Williamson and J. A. Greenwood, Tube Investments Ltd. (see *Proc. Roy. Soc.*, **237**, 560).
13. Pick-up of plastic lubricants in dry transfer—lubrication and its significance for sliding electric contacts.
    A. Fairweather and D. G. M. Shirley, Post Office Research Station, Dollis Hill.
14. Electrodeposition of the precious metals.
    J. R. Knight, Engelhard Industries Ltd., F. H. Reid, The International Nickel Co. (Mond) Ltd., and J. E. Philpott, Johnson, Matthey and Co. Ltd.
15. Practical aspects of light duty electrical contacts.
    J. C. Chaston, Johnson, Matthey and Co. Ltd.
16. Finishes for electrical contacts.
    R. W. Beattie and J. J. Miles, Automatic Telephone and Electric Co. Ltd.
17. Electroplated contacts—an evaluation.
    F. G. Swift and C. A. Abrahams, Standard Telephones and Cables Ltd.
18. Electrical contacts in aircraft switchgear.
    J. G. Steel, Central Electricity Generating Board.
19. Light duty contacts at elevated temperatures.
    W. Betteridge and H. C. Angus, The International Nickel Co. (Mond) Ltd.
20. The selection of electrical contact materials for light duty applications.
    V. G. Mooradian, Engelhard Industries Inc., H.A. Division, New Jersey, U.S.A.
21. Investigation into the electrical contact properties of granular carbon aggregates.
    E. D. Macklen, Standard Telecommunications Laboratories Ltd.
22. Some electrical properties of non-rectifying semiconducting contacts: Behaviour of granular aggregates of carbon and silicon carbide.
    A. Fairweather and D. G. M. Shirley, Post Office Research Station, Dollis Hill.
23. Reliability of bound connections.
    A. C. Delamare and R. T. Weeks, Standard Telephones and Cables Ltd.
24. Present researches in electrical contacts as used in telecommunication in Japan.
    Goro Matsumoto. Tuhuko University, Sendai, Japan.
25. The influence of alloy structure on the transfer of matter.
    W. Merl, Dr. E. Durrwachter-Doduco-K.G. Pforzheim, Germany.
26. Summary and closing remarks.
    M. R. Hopkins, University College of Swansea.

It is hoped that some five or six of these papers may be submitted for subsequent publication in this journal.

The physics of electrical contact phenomena

by Professor F. LLEWELLYN JONES, M.A., D.Phil., D.Sc., F.Inst.P., Department of Physics, University of Wales, University College of Swansea

Address given on 5th April 1961, at the opening of The Institute of Physics and The Physical Society Symposium on Electrical Contacts

MS. received 1st May 1961

Abstract

An outline is given of the nature of the fundamental physics processes occurring at an electrical contact and the problems to which they give rise, particularly in relation to light duty electrical contacts. The discussion includes contacts in which currents and potentials are of the order of amperes and volts, and also the so-called electrostatic contacts, in which one or other or both of these quantities may be extremely small. Recent work on microscopic molten metal bridges, micro-arcs (both of which are important in metal transfer) and the problems of 'electrostatic' contacts which mainly depend on surface properties, are described. Outstanding problems are discussed and the method by which they are being attacked are indicated.

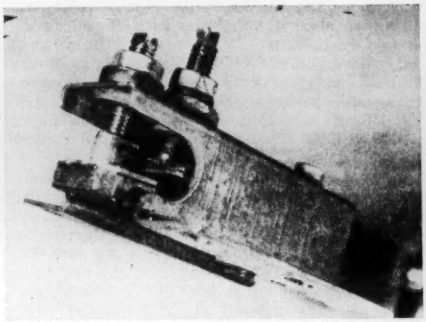
Introduction

THE problems of contacts, as might be expected, arose first in acute form in electrical technology and are familiar to electrical engineers, but in recent years it has become clear that the complete elucidation of the contact operation involves fundamental considerations in many branches of physics. In the various operations required of a contact: making, maintaining and breaking a circuit and then repeating this sequence at specified intervals, the physical conditions vary widely. It is therefore convenient to divide all contacts into three broad classes: heavy duty power contacts, e.g. the switches, contactors and circuit breakers of power installations; light duty communication contacts, e.g. relays in telephone circuits, contact breakers in ignition systems, etc., and contacts used in electronic apparatus—so-called electrostatic contacts. The problems of heavy duty contacts are mainly those of the extinction of powerful electrical gas discharges. These problems are regarded as being outside the scope of this Symposium.

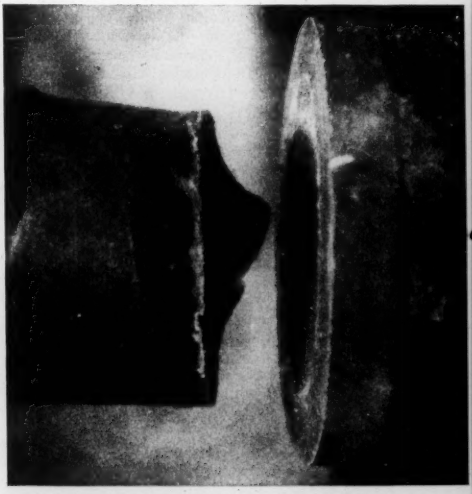
In light duty contacts the main problem is that of metal transfer between the electrodes and its effects. A typical example of such effects is illustrated in Fig. 1 which shows the well-known pip and crater formation.

In the so-called electrostatic contacts, the situation is dominated by the presence and properties of thin tarnish films which may lead to contact failure or at least to signal distortion.

The wide range of the problems involved and the various methods of investigation are well illustrated in the scope of the papers in this Symposium. In this opening address, my purpose is to discuss briefly the fundamental principles involved in the physics of electrical contacts.



(*a*) A typical vibrator contact assembly.



(*b*) Magnified picture of pair of contact electrodes from the above assembly showing pip and crater formation.

Fig. 1. Result of excessive fine transfer at a vibrator make-and-break contact.

Light duty contacts

In a survey of this kind it is helpful to analyse the processes which take place in the complete operation of an actual contact as they occur in chronological order. First, there is the approach of the contact electrodes as their separation becomes extremely small, say 10^{-4} cm. Secondly, there is the phase when contact is established and a steady current passes. This includes the heating of the contact and the formation of the molten metal bridge between the electrodes as they begin to separate. Finally there is the breaking of the contact, the arcing which accompanies it and the consequent damage to the electrode surfaces. This method of considering the subject, though not the simplest, helps to bring out not only the problems themselves but their interrelation.

The pre-contact arc

When two plane cold inactive electrodes are at a distance d apart in an atmosphere at pressure p, no breakdown and thus no discharge can occur if the potential difference V is less than V_s, the static breakdown potential given by the Paschen curve. When V exceeds V_s, under certain conditions, discharge can occur. If circuit conditions permit an arc 'follow-through' also takes place. V_s is a function of the product pd and of the nature of the gas. The initiation of breakdown requires the presence of initiatory electrons at the cathode or in the gap, as well as the lapse of sufficient time—the formative time lag—for ionization to develop. If there are no initiatory electrons a discharge will not occur. The shape of the Paschen curve is well known; it has a well defined minimum V_m of about 250–300 v. If V is less than V_m spark breakdown should not occur under the conditions specified above. (It is for this reason that the potential difference across contact breakers and contact makers is in general practice kept below this value.) Although this conclusion is consistent with experimental evidence at large distances (of the order of 1 cm) the situation is different in practice at very small gap distances. In fact, at small distances less than about 10^{-4} cm, it is well known that a discharge can take place even with gap potential differences of 10 v to 50 v in gases as well as *in vacuo*. Work in the Bell Telephone Laboratories in the U.S.A. has drawn attention to the importance of this phenomenon in relation to contact erosion. Fig. 2 illustrates the occurrence of an arc at the making of a contact as well as the better-known arc at the breaking of a contact.

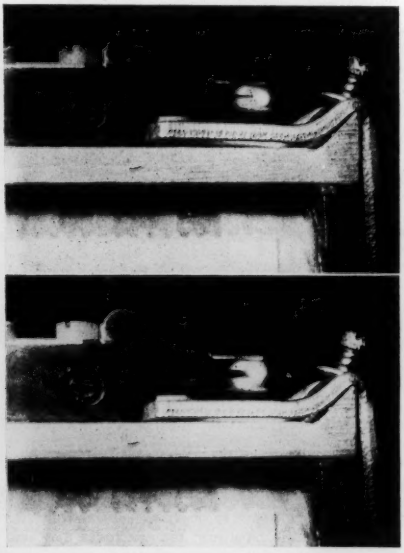


Fig. 2. Discharges at a contact.
Top: discharges occurring *before* the contacts touch.
Bottom: the arc discharge at break.

It is not easy to see how such a discharge can be initiated with cold inactive electrodes, but it becomes explicable if the cathode is electrically active as an emitter of electrons though still cold. It has been established that this 'cold emission' requires electric fields no greater than 10^5 v/cm provided that the surface of the electrode is covered by a thin tarnish film. Positive ions located on the film can play an important part in this surface activity. A detailed description and elucidation of this activity involves many aspects of surface physics as well as the physics of discharges. Thus thin films are of importance, not only in relation to the behaviour of electrostatic contacts, but also in the emission of electrons from contact electrode surfaces, which is a decisive factor in arc initiation.

Contact and the bridge

Proceeding now to the next stage of the contact operation, it is well known that due to the microscopic topography of the approaching surfaces, contact is first made over an increasing number of microscopically small areas. For a given circuit current, the current density at these small areas, though initially high, rapidly falls as the contact area increases until, when the contact between the electrode is fully made, negligible local heating is produced at the contact.

Consider next the opening of the contact in which these phenomena occur in reverse. The total area of contact falls, and the current density rises, until pronounced local heating occurs at the last microscopic area of contact. The constraining of the lines of flow of the current I in this way introduces a 'constriction resistance' R_c which sets up a potential difference ($=R_cI$) at the constriction. R_c is of the order of $\rho/2a$ where ρ is the specific resistance of the contact material and a the radius of the constriction area, assumed circular. Measurements of R_c show that a is of the order of 10^{-4} cm yielding current densities of the order of 10^7 A/cm^2.

Now, the relation between the bulk properties of the material and the observables, like contact potential difference V and the maximum temperature θ_m, has been the subject of much interest since the turn of the century. The relation between V, the maximum temperature θ_m and the room temperature θ_0 in a contact is given by the well-known equation which, on certain simplifying assumptions, can be expressed in its simplest form as

$$\theta_m^2 - \theta_0^2 = V^2/4L,$$

where L is the Lorenz constant ($=2{\cdot}45 \times 10^{-8}$ w ohm deg^{-2}). Thus taking θ_0 as about 300° K, and $V = 1{\cdot}5$ v, $\theta_m = 4600°$ K, which is above the boiling point of platinum. Thus a comparatively small potential difference across a contact can produce boiling of any metal. Because this result is basic

to the understanding of the temperature–voltage dependence in a contact assembly, this important relationship between potential difference and temperature has been put in its most general form (the (ψ, θ) relation), removing the simplifying assumption, and it is important to realize that the (ψ, θ) relation is independent of the geometry of the system, and that thermoelectric effects are also taken into account. It can be shown that heat lost by radiation from the surface of the contact bridge is negligible compared with the heat conducted through it, so that in the conducting system isothermal and equipotential surfaces coincide and the steady state relation then applies. If ψ is the generalized potential defined by

$$\mathbf{j} = -\kappa \operatorname{grad} \psi$$

where $\mathbf{j}$ is the current density and κ the electrical conductivity, then

$$\psi^2 = F^2 - 2\mu$$

where

$$F = \left(2\int_{\theta}^{\theta_m} \frac{\lambda}{\kappa} d\theta\right)^{1/2}$$

λ being the thermal conductivity, θ the temperature where the potential is ψ, θ_m the maximum temperature and μ a small term due to the Thomson effect. This relation is independent of the geometrical configuration of the system.

If S_1 and S_2 are two isothermal surfaces at temperature θ_0, one on either side of the bridge and comparatively remote from it, then the difference between the ψ's at these surfaces is the same as the difference between the ordinary electric potentials there. If V is this potential difference (i.e. the potential difference across the contact) then

$$\tfrac{1}{2}V = F(\theta_0, \theta_m) = \left(2\int_{\theta_0}^{\theta_m} \frac{\lambda}{\kappa} d\theta\right)^{1/2}$$

or

$$V^2 = 8\int_{\theta_0}^{\theta_m} \frac{\lambda}{\kappa} d\theta.$$

It is clear that this reduces to the simpler form given above in the particular case when the Weidemann–Franz law holds. The maximum temperature in a bridge of a given metal is a function of the potential difference across the contact. As V is increased θ_m is increased. Thus, as the electrodes of the contact are further withdrawn R_c increases, so that V increases and the maximum temperature in the bridge rises; the rise of temperature increases R_c further. Thus, the maximum temperatures can reach the melting point, and the mainly cold electrodes are thus joined by a microscopic globule of metal which forms a current carrying bridge between the two electrodes.

Metals of high melting points can thus form bridges which glow brightly, and this has been established experimentally. As the electrodes are separated further the bridge is drawn out, and its resistance consequently increased. The V is thus increased and θ_m consequently rises, eventually approaching the boiling point.

Stability and shapes of bridges

The microscopic molten metal bridge has been the subject of much attention and investigation on the part of physicists in recent years, for it is generally recognized to play an important part in subsequent phenomena during the complete opening of a contact. Perhaps it may be emphasized, at this point, that in normal practice the scale of these phenomena is very small. With a stable bridge the gap between the contact surfaces is usually about 10^{-4} cm, and the volume of the bridge is of the order of 10^{-10} cm^3.

The shape of a molten bridge is determined by surface tension considerations and the surface may be either a nodoid, an unduloid or a catenoid. These are the well-known surfaces of revolution whose profiles are the curves traced out by the focus of a hyperbola, an ellipse and a parabola, respectively, rolling on a line. In the case of a nodoid, the pressure in the liquid is less than the external atmospheric pressure, in the case of an unduloid (of which the cylinder is a special case), greater. The catenoid is the intermediate case in which the internal and external pressures are equal. These are the shapes that would be expected if the surface tension were uniform. Since surface tension varies with temperature, the equilibrium shape of a molten metal bridge must depend, in the case of a clean metal, on the viscous forces called into play by internal circulatory motion. Where there is a contaminating surface film, the beginning of the circulatory motion will redistribute the thickness of the film so as to produce uniform surface tension in spite of the variation of temperature.

The four photographs of Fig. 3 show the development of a molten quasi-static iron bridge.* These magnified pictures are those of a bridge which is much larger than those usually encountered in practice in contacts. The shapes conform to those which are predicted theoretically.

For the circuit to open, the molten metal bridge must break, and the way in which it does this determines some of the subsequent events. Considering, for example, a bridge formed equally from the metal of both electrodes, any asymmetry in rupture would provide a possible cause of metal transfer from one electrode to the other. Consequently, temperature asymmetry in the bridge has been, and still is, the subject of much investigation. In fact, the shift δ of the hottest section of a bridge from the position of symmetry due to the Thomson effect is given by

$$\frac{\delta}{AR} = \frac{\int_{\theta_m}^{\theta_0} \kappa dG}{\int_{\theta_m}^{\theta_0} \kappa dF}.$$

$2R$ is the 'reduced resistance' of the system, A the cross-sectional area of the wire near the section S which divides the reduced resistance into two halves and F and G are functions defined by

$$\psi = F - G \quad \text{on the right of S}$$
$$\psi = -F - G \quad \text{on the left of S,}$$

ψ being the generalized potential. G is a function of the Thomson coefficient.

Electronic effects due to the presence of thin insulating films have also been considered and examined experimentally as possible causes of metal transfer in contacts.

Again, when temperature asymmetry is important, the amount of metal transfer should clearly depend on the shape of the bridge before rupture, and the physical reasons for the ultimate shape which the bridge takes on when about to break is therefore also a subject of study.

Theoretical considerations on these lines indicate that the bridge actually visible between the electrodes may well be

* In the lecture at the Symposium a ciné film was shown, partly in black and white, partly in colour, illustrating the sequence of shapes assumed by such a bridge at various stages of its development.

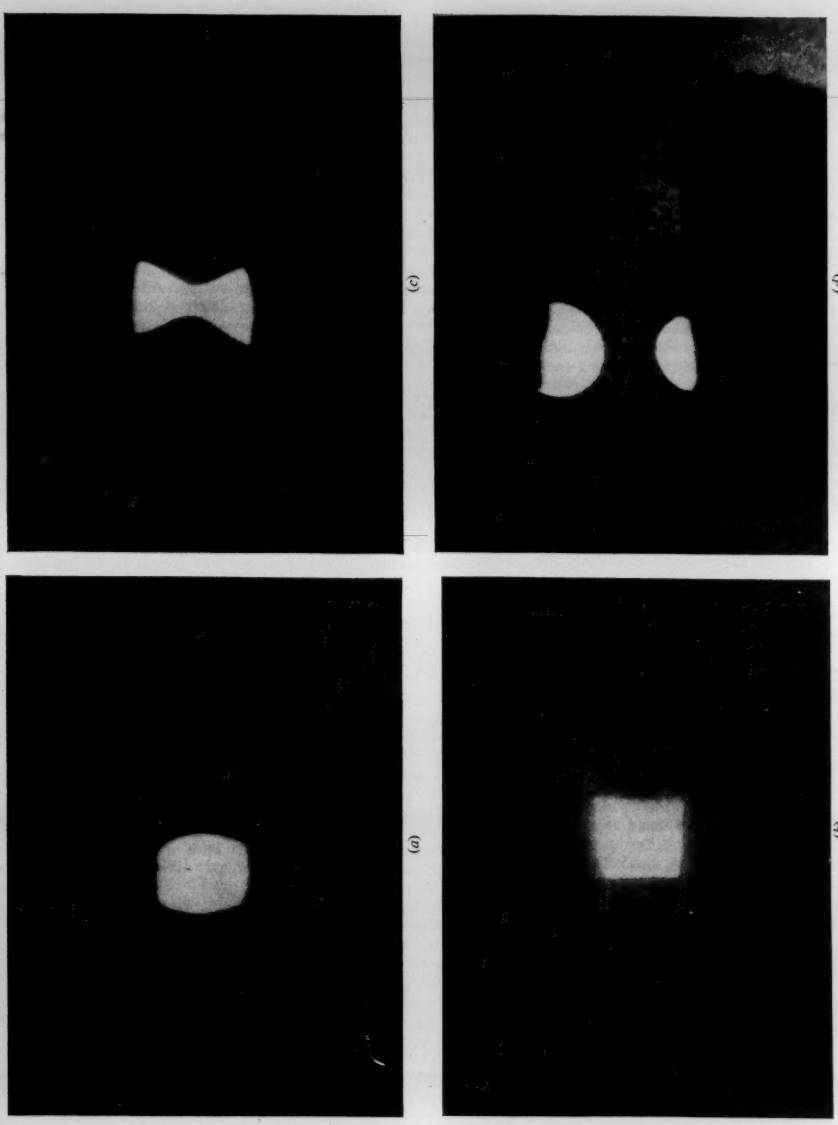


Fig. 3. Development of a molten iron bridge showing the sequence of shapes.
(*a*) Initially an unduloid. (*c*) Final nonoid.
(*b*) Catenoid. (*d*) The separate hot parts of the ruptured bridge.

only a part, in some cases a small part, of the complete bridge, the great part of which is inside the electrodes and invisible. This is one of the main reasons why, for many years, attempts to relate the amount of metal transferred in one operation to the supposed volume of the bridge met with little success. Firstly, the methods of measurement—gravimetric or optical—were inaccurate, and, secondly, the volume taken for the bridge itself was wrong.

Since the amount of metal transferred must depend on the ultimate size of the bridge as well as on its shape; the question of stability and of how the bridge breaks is of some importance.

Among numerous processes which can lead to instability might be mentioned the rather obvious ones of actual boiling when θ_m attains the boiling point, the variation of thermal and electrical properties with temperature (the (ψ, θ) relation shows how maximum temperature depends on bulk properties and voltage), electrodynamic forces 'pinching' the bridges, and surface tension forces.

Among the rather less obvious causes may be cited internal circulatory motions produced by viscous forces in the bridge, and elementary circuit relationships depending on the e.m.f. in the circuit.

In any case whatever the actual cause, or combinations of causes, the temperature of the bridge will suddenly rise and boiling will result.

In fact, careful measurements have shown that the rupture temperature is very near the boiling point and that in those cases examined in detail, actual boiling has occurred.

Break

The phenomena which occur immediately the bridge breaks—or blows up as it appears to do—is a matter of great importance in contact physics. It is practically at this instant that the metal transfer takes place, but the amount transferred in a single operation is exceedingly small, being 10^{-11} or 10^{-12} cm^3. Accurate measurement of such rates of transfer by optical or gravimetric methods has proved very difficult if not impossible, but application of radiotracer techniques in recent years has proved very encouraging, and reliable estimates of the amount and directions of metal transfer and its relation to the bridge can now be made. This method is about 1000 times more sensitive than the previous gravimetric methods, and has enabled the actual migration of the electrode material to be identified. As this work proceeds it is confidently expected to provide valuable information about the role which the bridge really plays, and possibly also help to elucidate the interesting processes occurring just after the bridge blows up.

Indeed, data so far obtained by this method at Swansea, in which the dependence of transfer on local inductance and the destiny of the matter from the exploded bridge were examined, are consistent with the setting-up of a micro-arc when the bridge explodes, and that migration occurs by ionic drift in the high pressure plasma formed. This brings us from considerations of the metal physics of the bridge back to fundamental considerations of plasma physics.

Post-break arcing

All this, however, is still not the end of the complete contact operation.

Consideration of the state of affairs immediately the bridge breaks shows that the fundamental conditions for the initiation and maintenance of an arc discharge are simultaneously satisfied. These are thermionic electron emission from the cathode because of the high local temperature, the presence of a gas atmosphere of high density and low ionization potential (metal vapour), and occurrence of a high potential difference between the electrodes; an induced e.m.f. set up in the presence of self-inductance determines the magnitude and sign of the e.m.f. when the current falls at 'break'.

This latter point is very important because as soon as this is recognized the dependence of metal transfer on the local small inductance of the circuit (down to 10^{-8} H) becomes an important matter of enquiry.

Many other problems on the arc naturally follow. Gas pressure–gap voltage relationships are particularly significant as they can influence the net transfer—whether it is due to ionic or electronic bombardment of the electrodes, for instance.

Electrostatic contacts: thin films

Finally, we come to electrostatic contacts in which the role of thin surface film is of paramount importance.

When the contact potential difference is as low as microvolt and the current as low as microampere, the phenomena so far discussed do not occur.

Except when the electrodes have been specially cleaned and maintained in a vacuum or in a pure rare gas, the electrodes of a contact in ordinary engineering use are exposed to the deleterious effects of corrosive or tarnishing atmospheres. Contamination may consist of surface films produced by adsorption, tarnishing or erosion, as well as the presence of oil, grease or dust particles. Films of oxides, carbonates or sulphates are likely to form in industrial districts.

When the contact voltage is low, the heat produced by very small currents is too small to decompose or drive off such films; further, in these applications, the electric field set up across them may be insufficient to break them down. Such films are not only bad conductors or good insulators, they are often semiconductors, so that when the signals themselves are not actually interrupted, they may well be distorted. Considerable investigation of corroded contacts and the mechanical properties of their surface films have been made at the Research Laboratories of the British Post Office, where the part played by these films in sliding contacts, and particularly the influence of PTFE as a lubricant in such contacts have been studied.

Lastly, another important contact problem arises in relation to carbon brushes and the mechanism of abnormal wear. Interesting work in this field has been proceeding in the Research Laboratories of General Electric at Schenectady, U.S.A.

Conclusion

In conclusion, I would say that I have tried to survey the whole vast field of the fundamental physics of electrical contacts, to point out some outstanding problems and to indicate how these are being tackled today.

Hydrodynamic research

by F. S. BURT, B.Sc.(Eng.), A.C.G.I., D.I.C., F.R.Ae.S., Admiralty Research Laboratory, Teddington, Middlesex

Discourse given at the Annual Exhibition of The Institute of Physics and The Physical Society, 17th January 1961

Abstract

A brief survey is given of some hydrodynamic research problems of particular interest to naval and marine applications. A description is given of model ship towing tanks and their use in studying resistance and propulsion problems of ships. The use of special sea-keeping and manœuvring basins from model ship studies of sea-keeping and manœuvring characteristics is mentioned. Another topic is the study of ship propulsion research and a description is given of cavitation tunnels and their use in this connection.

Some detail is given of the particular sphere of underwater hydrodynamic research and the special facilities and instrumentation which have been developed at the Admiralty Research Laboratory for research into this specialized field including the use of slotted wall working sections in water tunnels and the large rotating beam channel.

The unusual hydrodynamic research problems associated with the entry of missiles from the air into the water are briefly surveyed as are those of two-phase flows in gas liquid mixtures.

The discourse concludes with a brief mention of possible high performance vessels of the future including hydrofoil craft, hovercraft and underwater cargo vessels.

Introduction

HYDRODYNAMIC research is a very broad field, so the scope of the discourse is limited to aspects of hydrodynamic research of particular interest to naval and marine applications. Most emphasis is given to experimental hydrodynamics research techniques and facilities.

As in many cases full-scale experimentation is slow and costly and can rarely be carried out under controlled conditions, most of the experimental work is done using models in installations designed especially for hydrodynamic research and often costly and of large size. The technique of scaling up from models to full scale is well established although there are many scaling problems not yet fully understood.

Resistance

In surface ships of displacement type, the hull resistance consists of two predominant components, (i) skin friction and (ii) wave resistance. Unfortunately, the two are governed by different laws, the one corresponding to Reynolds number and the other to Froude number, which makes it in most cases virtually impossible to simulate full-scale conditions correctly in model experiments.

The standard procedure used for measuring the resistance of model ship hulls is to satisfy the dynamic law, producing the proper wave pattern and to correct the skin frictional component on a basis of Reynolds number.

In larger displacement vessels skin friction along hull predominates, for example in transatlantic liners friction drag is about 75% of the total, whereas in a small high-speed motor boat wave drag predominates.

Wave resistance

Corresponding to the pressure increase which occurs at the bow of a hull, the water level rises in the vicinity of the stem. On the other hand at the sides of the hull, particularly in the vicinity of the shoulders, the static pressure is decreased. Hence a wave system from a ship's hull originates from a crest at the bow and a similar but somewhat weaker system originating from a trough at the stern. The dissipation of energy in this wave-making results in a resistance to the ship's motion. As far as the hull is concerned wave drag appears in the form of a net longitudinal component of the pressure distribution along the hull.

The theory of the drag due to wave-making has been developed by many scientists, notably in this country by Professor Havelock. It has been demonstrated that on the average the transverse and the divergent part of a wave system each roughly contributes half of the wave drag. The transverse waves, however, determine the interference between the bow and stern systems. Depending on whether these two systems are in phase or not they can add to each other or they may partially cancel each other. As a consequence, the drag coefficient of a displacement craft shows 'humps' and 'hollows' as a function of speed.

Model ship towing tanks

Since the theory of wave resistance is complex its magnitude is customarily derived from model tests in towing tanks. Alterations in hull shape may be considered in terms of wave-making length. By appropriate design it is possible to make the crusing speed of a ship coincide with a hollow in the wavemaking resistance curve rather than a hump.

There are now a large number of ship towing tanks around the world for studying resistance and also propulsion problems of ships. The earliest one was William Froude's original Admiralty Experiment Works at Torquay where experiments in the study of ship resistance were conducted from 1871–1885. The original tank was then replaced by a longer tank, the No. 1 Ship Tank at the Admiralty Experiment Works, Haslar, which has been in continuous use since

it was completed in 1887. It is 400 ft long, 20 ft wide with a maximum depth of 9 ft. This tank was supplemented in 1932 by the No. 2 Ship Tank, which is 890 ft long, 40 ft wide and 18 ft deep.

As in most ship tanks the model is carried below a special towing carriage which spans the tank and is driven by four electric motors, one at each corner, along locomotive type rails set most accurately on the walls on either side of the tank. This carriage has a designed maximum speed of 40 ft/s and can maintain uniformity of set carriage speed to within $\frac{1}{4}\%$.

A feature of these towing tanks is the necessity for rapid acceleration and braking so as to provide the maximum length run at steady speed.

One of the most recent large towing tanks is that built for the Department of Scientific and Industrial Research in the new Ship Hydrodynamics Laboratory at Feltham.

The towing carriage for this tank which is some 50 ft × 50 ft in plan, is built of tubular steel members and weighs about 40 tons. It is driven by four 300 h.p. motors, one at each corner, at speeds of up to 50 ft/s.

The ship models in these tanks are connected to the towing carriage via a special balance system on which the forces on the model can be measured. In most of these dynamometers, the major portion of the forces being measured are balanced by weights and the residue only is measured by spring balance.

To obtain the best shape of hull lines and the best shape and position of the appendages, towing tests are first made with the bare hull and its resistance characteristics measured over a speed range, the model is then tested with the appendages added and finally with the propellers fitted in a self-propelled version.

Sea-keeping

The standard towing tank tests refer only to calm water conditions, whereas in real life, unfortunately, the sea is rarely completely calm and we are interested in the behaviour of ships in a seaway. Consequently, some of the towing basins are fitted with wave-making facilities. These tests in waves are an essential part of research into seaworthiness, as the shape which produces the lowest resistance characteristics in calm water is by no means necessarily the optimum form from seaworthiness considerations.

The wave-makers fitted to towing tanks are limited to simple wave systems which travel down the length of the tank and therefore strike the model exactly on either the bow or the stern, which is a rather idealized condition compared to what happens in the complicated wave patterns which occur at sea. The wave-maker in the No. 2 Ship Tank at Haslar is typical and can generate waves of up to a wavelength of 40 ft and a height of 2 ft, thus reproducing on a model scale the full range of ocean waves.

As the importance of ship behaviour in rough seas is now more fully appreciated there is an increasing emphasis in research on the sea-going qualities of ships. New experimental facilities have been and are being built, and there have been notable advances in ship motion theory and an increasing amount of full-scale work is being done. It is certain that as in many other fluid dynamic research fields experimental work on models is most valuable when it is supplemented by both theoretical and full-scale studies.

Amongst the new facilities in the world for sea-keeping research there are several capable of oblique wave tests. Most of these facilities take the form of square or rectangular basins fitted with wave-makers on two adjacent sides enabling oblique and confused seas to be generated. A particularly fine facility in the form of a 200 ft × 400 ft basin fitted with plunger type wave-makers on two sides will shortly be in full operation at Haslar.

A most unusual and versatile sea-keeping laboratory is that of the Netherlands Ship Model Basin at Wageningen. The basin is 100 metres long and $24\frac{1}{2}$ metres wide and $2\frac{1}{2}$ metres deep and is fitted with pneumatic wave-makers along one side and the end, with wave absorbers in the opposite sides. These pneumatic wave-makers are electrically programmed to produce irregular wave patterns. The model is held with the minimum restraint below a towing carriage which travels on a pair of rails mounted on pillars set in the centre of the tank.

Manœuvring

In the past the design and development of surface craft has been basically concerned with improvements aimed at obtaining low drag forms. The problems of manœuvrability and controlability until the last decade were never considered as major design criteria.

A technique being used to obtain full-scale data on ship handling is the measurement of data obtained during spiral and zigzag manœuvres. The tools and techniques for the determination of spiral and zigzag manœuvres from model studies are at present in existence. Free running self-propelled models equipped with the appropriate instrumentation are run in manœuvring basins either under radio control or programmed control. A typical facility is the 100 ft square sea-keeping and manœuvring tank at Ship Hydrodynamics Laboratory, Feltham. The data is either recorded in the model or ashore and the course of the model plotted either by analysis of photographs or in the case of the Ship Hydrodynamics Laboratory tank a special tracking system is being devised which will enable the course to be plotted automatically, thus saving a great deal of analysis.

Detailed studies of the hydrodynamic factors from which many forms of trajectories can be calculated can only be undertaken by captive model studies in curvilinear motion. A facility used for this type of work is the rotating arm basin of which there are now quite a number in the world.

Propulsion

The objective of propulsion research for ship design is to obtain the best shape of hull and layout of appendages plus the optimum design of propeller, both for maximum full speed and for maximum economy at the design cruising speed. These two requirements may well conflict and a compromise has to be reached.

In research studies to investigate propulsion by means of models, three principal types of test are used.

(i) Open water test carried out in a model basin with the model screw propeller ahead of a special test rig and dynamometer.

(ii) Self-propulsion tests with a model in a ship tank, carried out in Britain by the constant speed method, in which experiments are made which cover a whole range of propeller loadings at several set speeds.

(iii) Tests of model propellers in a cavitation tunnel for determining the occurrence of cavitation and the behaviour of the propeller under cavitating conditions.

Cavitation is a dynamic process in which cavities filled with vapour are formed in the fluid in regions where the

local pressure falls to that of the vapour pressure of the fluid. On ship propellers cavitation makes a noise, may cause erosion and also may set up vibration, so every effort is made to minimize it.

The experimental facility used for research into propeller design and its cavitation performance is the cavitation tunnel.

All water or cavitation tunnels consist essentially of:

(i) a working section in which the model is mounted and can be observed;

(ii) a closed water circuit consisting essentially of a high capacity pump and piping by means of which the flow of water through the working section is maintained;

(iii) a control system which enables the water velocity and pressure in the working section to be regulated over a wide range of operating conditions;

(iv) a system of instrumentation enabling the relative parameters to be measured, in any given experiment; and

(v) in a propeller cavitation tunnel, a separate electric motor and drive to the propeller being tested in the working section fitted with dynamometer equipment enabling thrust and torque to be measured.

The latest tunnel of this type built in this country is the new 44 in. diameter cavitation tunnel which is part of the new research facilities at Ship Hydrodynamics Laboratory, Feltham. In this tunnel propellers of up to 24 in. diameter can be tested and are driven by a 300 h.p. motor. The pressure in the test section can be varied from near zero to six atmospheres absolute.

The normal experimental procedure is to vary the propeller rev/min at constant speed and pressure in the working section so that a range of loadings is covered. This is done for a number of speeds and pressures and at the same time observations are made of any cavitation phenomena, by using stroboscopic lighting, and the cavitation can be photographed if required.

Cavitation

The theory of cavitation erosion is complex and there are theories of mechanical, mechanochemical, electrochemical and thermoelectrical damaging due to cavitation. One of the most tenable theories of purely mechanical damaging is that the explosion of cavitation bubbles cause fatigue and resultant damage of the surface. Once mechanical damage has taken place corrosion may ensue.

Non-uniform wake

A cavitation tunnel produces a uniform velocity stream, whereas in fact propellers when fitted to a ship operate in a non-uniform flow as they are working in the boundary layer of the ship's hull and in the wakes of any appendages. This naturally has a most marked influence on the performance of the propeller.

Various techniques are used to obtain the full-scale wake in the propeller tunnel. The No. 2 tunnel at Haslar is large enough for the propeller to be tested behind the appropriate portion of hull of the ship mounted up-stream of the propeller. Another technique is the use of wire meshes up-stream of the propeller to produce a copy of the full-scale non-uniform wake.

Another interesting approach is that used by the No. 2 Water Tunnel at Netherlands Ship Model Basin, Wageningen. This tunnel has a special flow regulator fitted in the contraction ahead of the working section. This divides the whole velocity field into a number of separate elements, the flow through each being adjusted by a needle valve. In such a regulator only the axial components can be varied but these are probably the most significant.

Research is also being carried out into alternative systems of propulsion to standard propellers; these include Voith–Schneider propellers; and also into various forms of jet propulsion somewhat analogous to gas turbine propulsion in air.

Underwater hydrodynamics

A different sphere of hydrodynamics is that concerned with the study of the characteristics of bodies which travel under the water. This research is directed towards the better understanding of the science and the resulting improved performance for future underwater vehicles whether they be weapons or submarines, and is the particular branch of the science which is the concern of the Fluid Dynamics division of the Admiralty Research Laboratory. Underwater hydrodynamics research bears a close resemblance to the sister field of aerodynamic research. In fact some of the earlier work on drag reduction and stability and control of airship forms which took place in the 1920's is extremely relevant.

To carry out experimental work in this field suitable experimental equipment has to be provided and at the Admiralty Research Laboratory an impressive and unique set of facilities exist for doing this work. The main equipment comprises a large rotating beam channel and two water tunnels, in addition there is a glass-sided tank used for investigations of the problems of the entry of missiles from air into the water.

In the aeronautical research world an essential tool is the wind tunnel and the two Admiralty Research Laboratory water tunnels are in fact very similar to wind tunnels in their concept and differ somewhat in layout from the ship propeller cavitation tunnels referred to earlier.

The smaller of the two tunnels has a 12 in. diameter working section and is very versatile in that it can be fitted with three alternative working sections:

(i) a conventional closed jet 3 diameters long, (ii) an open jet $1\frac{1}{4}$ diameters long, and (iii) a slotted wall working section 6 diameters long.

The water speed in this tunnel can be continuously varied up to some 90 ft/s and the pressure in the working section varied between near zero and 3 atmospheres absolute.

Slotted wall working section

A most important part of this research is naturally concerned with powered underwater bodies and this imposes quite a problem in the design of the facilities. The smallest possible size one can make a satisfactory powered model of a torpedo, say, with representative scale propellers, is about 8 in. in diameter. In the conventional closed jet section one cannot normally test models more than $\frac{1}{8}$ of the diameter owing to excessive blockage corrections, which would mean a working section of at least 48 in. diameter. In addition most of the work is on quite long bodies and in a conventional closed jet the build up of the boundary layer on the tunnel walls results in a continuous change of static pressure down the length of the working section.

To overcome this impasse the concept of a slotted wall working section was developed (see Fig. 1). The slotted wall tunnel is well known as a tool for studying transonic flow. The tunnel choking and interference effects associated with

a closed tunnel severely restrict the size of model which can be tested at transonic speeds; on the other hand, open jet tunnels are unsteady, the jet length limited and it tends to break up. The slotted wall working section is a compromise, the solid part of the boundary steadying the flow and the

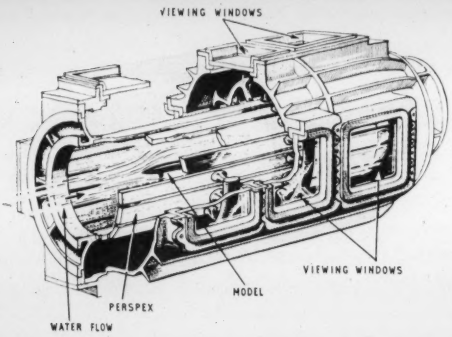


Fig. 1. Slotted wall working section for water tunnel.

open part of the boundary preventing choking and reducing interference effects. These advantages are still valid at subsonic speeds.

By development of these principles it has been possible to develop a slotted wall working section for water tunnels enabling models of up to $\frac{1}{4}$ of the jet diameter to be tested. In addition there is constant static pressure down the length of the working section.

30 in. water tunnel

The larger 30 in. diameter water tunnel (Fig. 2) was designed from the outset around a slotted wall working

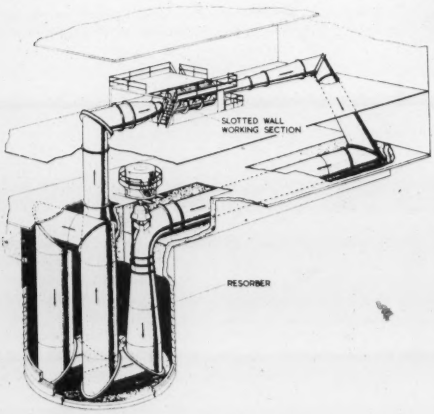


Fig. 2. General view of Admiralty Research Laboratory 30 in. water tunnel.

section to permit the testing of powered torpedo-like bodies of up to at least 10 in. in diameter and up to 10 ft long. This tunnel is also fitted with a 'resorber' which consists in this case of four vertical passes in a 60 ft deep pit. The principle of the 'resorber' is very simple. In any water tunnel when cavitating conditions are reached in the working section, vapour bubbles are formed and it has been found that they can be re-absorbed into the tunnel water as the pressure of the circuit is raised and the tunnel is kept running at this high pressure for a few minutes. The 'resorber' does this in effect automatically.

Rotating beam channel

Problems such as drag reduction, measurement of lift forces and turning moments, and the effect of depth on underwater bodies, are all best studied in a water tunnel where the model is stationary and the water flows past it. The water pressure on the model can also be varied over a wide range. Other vital problems, such as the study of the stability and control characteristics, and the investigation of underwater bodies travelling at speeds of the order of 100 knots are best studied in a rotating beam channel.

The 122 ft span rotating beam channel at the Admiralty Research Laboratory (Fig. 3) was specifically designed for

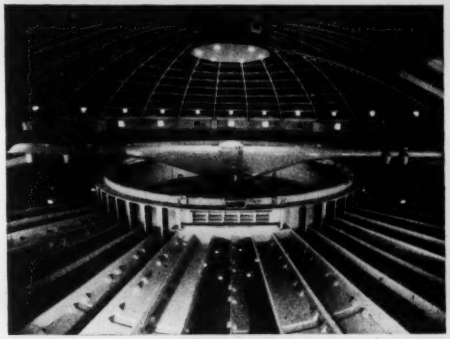


Fig. 3. The Admiralty Research Laboratory 122 ft span rotating beam channel.

this class of work. The 60 ton beam spins over a ring channel of rectangular cross section 34 ft wide and 15 ft deep. Models up to 20 ft long may be tested in this channel and are normally run at 5 ft deep. In stability and control problems one is concerned with the motion on curved paths as well as straight ones, and hence the need to measure the forces on models under these conditions. Models can be run at varying radii between 45 ft and 55 ft and the maximum speed of a model is 100 knots. In some studies of high-speed underwater bodies the circular path followed by the model is a nuisance and to avoid having yaw on the model in these cases it is necessary to curve the model to the running radius. The channel is also fitted with anti-swirl baffles and wave damping beaches which can be seen in the photograph.

So that the models may be observed under test, large underwater windows are provided at one section of the circuit in both the inner and outer walls of the channel. The million gallons of water in the channel are specially

filtered and treated to enable high-speed photographs to be taken of the cavitation phenomena which can be observed visually.

Instrumentation

The models used both in the channel and in the 30 in. water tunnel are in many cases self-powered and conveniently, for much of the work, the same 10 in. diameter model can be used either in the channel or the water tunnel. The models used are always fitted internally with any necessary instrumentation, for example a five-component balance measuring drag, lift, pitching couple, side force and yawing couple, pressures gauges to give pressure distribution around the body and in the case of a powered model, measurements of torque and thrust on the propeller shafts. The necessity for compactness and for electrical signal transmission in the case of the rotating beam led to the choice of resistance strain gauges as the basic measuring element, used in an a.c. Wheatstone bridge circuit with self-balancing potentiometers.

In addition a non-mechanical technique has recently been developed at the Admiralty Research Laboratory for continuously cleaning hot wire probes used for measuring turbulence in water. A d.c. cleaning current is passed through the water from the hot wire probe to a suitably placed electrode. The bubbles of gas which are then liberated at the surface of the hot wire are effective in removing particles of dust which adhere to the wire.

Water entry research

A major research problem with the many underwater weapons which start in the air and then have to travel under water is concerned with their water entry characteristics. To study these problems there is a very large glass-sided tank at the Admiralty Hydroballistics Research Establishment at Glen Fruin. This main tank is 150 ft long, 30 ft wide and 40 ft deep and the whole of one side is fitted with toughened glass panels enabling the water entry behaviour of the projectiles to be studied. In addition there is a similar but smaller tank 29 ft long, 9 ft deep and 5 ft wide at Teddington. Both of these tanks are fitted with slotted cylinder catapults for firing the models into the tanks.

These catapults have the advantage over a smooth bore gun in that they allow models to be fired with pitch and yaw to the line of the actual water entry shot. These conditions are those that tend to occur when missiles are released from aircraft.

The Glen Fruin catapult can fire models of up to 8 in. diameter at entry speeds of up to 500 ft/s, and that at Teddington models of up to 4 in. diameter at entry speeds of 300 ft/s. The model entry speed is scaled by the dynamic law, i.e. Froude number, so the model entry speeds of standard scale models at the two establishments, 5 in. and 2 in. diameter respectively, both correspond to 1000 ft/s full scale entry speed.

On the smaller models mirrors are mounted on the tail and an optical system is used to measure the angular orientation of the model whilst in flight. The larger diameter models used at Glen Fruin are fitted internally with transducers from which the signals are transmitted back via a trailing cable to an amplifier and recording system.

Most of the required data in water entry studies is recorded photographically using high-speed cine cameras. A typical still of a water entry photograph is shown in Fig. 4. The subsequent analysis of these photographic records presents a formidable problem. To overcome this analysis bottleneck a special film evaluator has been developed in which frame by frame the data on the film is punched out on standard five-hole teleprinter tape. This tape is then fed into a digital computer and the necessary computations are performed. The final data can either be printed out in a tabular form on a teleprinter or alternatively can be plotted in a graphical form on a plotting table.

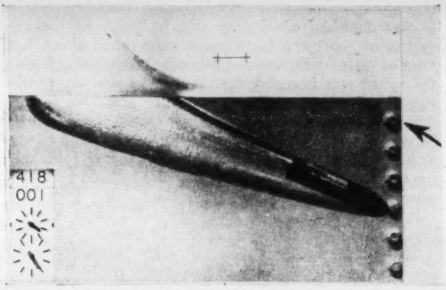


Fig. 4. A typical water entry photograph.

Gas–water mixtures

An interesting item of fluid dynamic research pursued at the Admiralty Research Laboratory has been a theoretical and experimental investigation of the properties of gas–liquid mixtures. This piece of fundamental research was started because gas–liquid mixtures were expected to have interesting properties and possible applications in a number of fields, notably to drag reduction, underwater propulsion and shock attenuation.

The first stage was purely experimental. A special tunnel for the production of gas–liquid mixtures was produced. Measurements have been made of the pressure loss due to friction in pipe flow of a mixture and also of the speed of propogation of weak compression waves through the mixture.

A theory of the propogation of plane shock waves through such mixtures has been developed and the theoretical results confirmed experimentally using a gas–liquid mixture shock tube. In this arrangement mixtures of good quality can be produced by bubbling air through glycerol in a vertical tube. Shocks can be propogated into the mixture by bursting a diaphram which closes the end of the tube. The passage of the shock wave is detected using a light source and a photoelectric cell.

Possible future developments

The two main components of surface ship resistance are wave-making resistance and skin friction resistance. If higher ship speeds are required something must be done to minimize the drag penalty. One possibility and the earliest considered historically is to raise the hull above the surface by using dynamic lift and this leads to the hydrofoil craft which has been under sporadic development since the beginning of the century.

In comparison with conventional fast boats, hydrofoil craft have much less resistance at high speed and the retarding effect of waves is greatly reduced as long as they are less

than the design height and sea craft should be able to operate in much worse sea conditions than conventional craft of the same size.

A second possibility is to go under the water; when a submarine runs far below the surface, surface wave-making resistance is virtually eliminated, and the remaining resistance is caused almost entirely by skin friction. By making a submarine short and fat, fineness ratio round about 7, the ratio of surface area to volume can be minimized. Even so, the wetted area is greater than that of an equivalent surface ship. This means at lower speeds the resistance is higher than a surface ship but at higher speeds the added skin friction is more than compensated by the reduction of wave-making resistance. Another great advantage of going under the surface is avoidance of the penalties suffered by the ship when it has to operate in rough weather.

It has been suggested that these advantages might be exploited in a submarine cargo ship or tanker.

The circular section submarine is best hydrodynamically and can show an advantage in horsepower required at speeds of say 24 knots but it would have a very large draught and it might be necessary to have an elliptical cross section to restrict draught to that of the surface ship, in which case the submarine would be at a disadvantage compared with the surface tanker in power required, owing to the greatly increased wetted area.

A recent development has been the concept of the ground effect machine or Hovercraft. This sprung from the idea of greatly reducing skin friction resistance by supporting the craft on a cushion of air. The principle of lift used by the hovercraft is the generation of an air cushion under the craft at a higher pressure than ambient, which is retained between the hull bottom and the sea surface by a peripheral air curtain. This principle is potentially of particular interest in the speed range 60–100 knots. There are, however, difficulties in operating over oceans with large waves owing to the limited clearance of the craft over the sea. Nevertheless, there are potential fields of applications particularly over sheltered waters.

Concluding remarks

It has only been possible to cover superficially a rather limited part of hydrodynamic research generally, but an attempt has been made to give some idea of the experimental problems and the research techniques used in this branch of science. It is a very fascinating subject and attractive experimentally. After all, most people like playing with water and this is doing it on a grand scale.

Physics of the ocean

by G. E. R. DEACON, C.B.E., F.R.S., National Institute of Oceanography, Wormley, Godalming, Surrey

Discourse given at the Annual Exhibition of The Institute of Physics and The Physical Society, 18th January 1961

Abstract

A hundred years ago better navigational instruments, improved charts and the compilation of wind and current atlases brought much improvement to navigation. In comparison the advances made by modern instruments and theoretical methods are second-order steps, but they are just as exciting, and with the increased pressure of to-day's civilization and industry, probably more important. In particular we are beginning to get a better understanding of waves, currents and the general circulation in the oceans, which will allow us to make better use of the oceans, and to contend with them on more equal terms when necessary.

THE first half of the nineteenth century was a stirring time for the science of the sea. After the war with France our sea trade doubled every fifteen years. Navigators had more efficient instruments, better astronomical tables, sounder nautical directions and better charts. Admiral Smyth, who used the post-war years to survey the Mediterranean Sea and soon became Vice-President of the Royal Society, President of the Royal Astronomical Society and President of the Royal Geographical Society, said in 1852 "The various branches of available science have been so steadily advancing among seamen of all nations that besides a higher practice of navigation they possess a more accurate information respecting the phenomena of winds and currents than heretofore." It was a remarkable all-round effort. Captain James Rennell, Surveyor to the East India Company, had made a very detailed study of the currents of the Atlantic Ocean, and Lieutenant Maury of the U.S. Navy, with a good deal of encouragement from this country, persuaded his Government to call the first International Maritime Congress in 1853. The outcome was a scheme by which observations of wind and current were collected systematically, and selected ships have completed meteorological logs ever since. Everyone was convinced of the importance of determining the best tracks for ships to follow. Admiral Fitzroy, first Director of the growing Meteorological Department of the Board of Trade, said in 1855 "During late years the great increase (by wider diffusion) of nautical knowledge has not only much shortened sea passages but has rendered them more secure."

A new approach does not, however, convince everyone. One of the first shipmasters to submit a meteorological log said "Remarks not so full or complete as I could have wished. It has entirely devolved upon myself, as I cannot get my officers to take any interest in the matter, being members of what is called the old school, they cannot or will not see the utility of bothering themselves (as they term it) with these new affairs." Although the powerful steamships of the later part of the century did not have to worry so much about wind and currents the work was continued, but it did not spread as much as might have been expected into new applications of science. The careful observations sent in by sailors over the past 100 years have made the charts of average wind and current published by meteorological offices all over the world about as good as charts of average conditions can be, and improvements in navigational instruments, tables, charts and sailing directions have been continued in a remarkable way, but there are many aspects of marine science which have received too little attention. Anyone seeking to further some aspect of navigation, fisheries, coastal engineering and marine meteorology, by applying a knowledge of waves, currents, deep water circulation and energy exchanges between the oceans and atmosphere soon finds the available information quite inadequate for his purposes. Disasters such as the sinking of the *Titanic* and decline of important fisheries have launched scientific expeditions and founded new laboratories, but it was only under the threat of something really overwhelming like the last war that there was real concentration of practical and theoretical effort on the science of the sea.

Our understanding of ocean waves affords a typical example. As early as 1876 G. G. Stokes realized that the alternation of groups of high and low waves which often occurs with enough regularity to allow the coxswain of a boat to take advantage of quiet periods is comparable with the beating of two sound waves, and due to interaction between trains of waves of different lengths. But it was not until 1944 that a group of Admiralty scientists made the first spectrum analysis of sea waves. It was then found that the waves behaved roughly as though they originated in a number of instantaneous point disturbances over a wave-generating area, each producing a continuous spectrum of waves so that the wave pattern at any time is the sum of a large number of component wave trains with random phase relationships. After they leave the storm area the component waves appear to travel independently with the group velocities appropriate to their periods, and if the generating area is well defined and far enough from the recording station the component wave trains can be distinguished one after the other by successive frequency analysis of the records. Under ideal conditions, as for example when the waves from a small but intense storm some 1500 miles distant arrive at the recording station during a period of calm weather, the slow decrease in the wavelength and period over two or three days is quite apparent even to visual observation.

The most noticeable change in waves entering shallow water is an increase in height. In deep water, deeper than half a wavelength, the speed of a wave depends on its length and is practically independent of depth, but in shallow water it depends on the depth of water and decreases as the depth decreases. As Rayleigh pointed out in 1911, its height then increases because it must carry energy towards the shore at more or less the same rate. The first swell to arrive from a distant storm may have a wavelength of as much as 3000 feet

and be only a few inches high so that it is not visible in deep water. It may, however, become apparent when it is shortened, heightened and steepened by passing over shallow ground and it is such modification of long forerunners from a distant storm that produces 'ground swell'. Most of the storm energy is usually contained in wavelengths about a third of the way down the spectrum and the arrival of the 'ground swell' is generally a warning of heavier swell to come.

There was no reasonably satisfactory apparatus for making a continuous record of waves before 1944. The most usual way to-day is to lay a pressure recorder on the bottom in soundings up to about 100 feet and to connect it to the shore by a conducting cable. The most advanced devices use the piezoelectric effect and, more recently still, frequency-modulated capacitance systems which are not so dependent on the state of the cable. All the pressure recorders have the fundamental disadvantage that they are not sensitive to the short wavelengths since the motion and pressure variations below a train of waves are reduced to about one-ninth at a depth of half a wavelength. Where there is a suitable mast or other fixed structure in the sea a vertical insulated wire can be used for measurement of the varying capacitance between the wire and the water as the water rises and falls. Since 1953 we have had a shipborne wave recorder which adds the variations of pressure measured through a small hole in the ship's hull below the water line to a record of the up and down movement of the hole obtained by doubly integrating vertical accelerations measured at the same point.

Recent work on the statistics of the moving wave surface is a remarkable scientific achievement, allowing precise treatment of the geometry of a very irregular surface which can be described in no other way. It gives useful information about the probability of different wave heights, slopes and accelerations, from relatively few measurements of a wave record. As an elementary example, one wave in 23 is likely to be more than twice the mean wave height, and one in 1175 more than three times the average. This is the unexpected wave that just happens when a large number of the component wave trains get into step. Statistical expressions of the oscillatory motions and stresses in ships can be calculated from the wave information and ship characteristics. It is generally admitted that the more precise description of the sea surface obtained by oceanographers has brought research on the design and sea-keeping qualities of ships to the threshold of new and exciting developments. It was the spectrum analysis of waves and ship motion by scientists making a systematic approach to the study of waves that provided the initiative. Naval architects and engineers primarily interested in ships might not have done it for a very long time. They have, in fact, been known to say that they were interested in the way ships moved, not in what makes them move.

There is still a great deal to do before we have real understanding of even the physical processes by which waves are generated. Two kinds of mechanism are now believed to be the most important: the effect of travelling pressure disturbances associated with the turbulence in the wind, and the growth of the interfacial disturbance as an instability in the shearing flow of the air over water. The first mechanism would produce an angular distribution of energy such that waves of particular lengths would be preferentially energized in the directions, on either side of the wind, where the component of wind velocity is the same as the wave speed. This means that the longer wavelengths would more closely follow the wind. The second effect might narrow the spread of all frequencies though the shorter waves are likely to be spread out again by impulsive effects such as wave breaking. There is still a good deal to learn about the propagation and decay of waves as they travel through other wave patterns and into changing depths and currents. We also need much more detailed study of the small transports of water by waves. Experiments and theory show that under certain conditions there is a rise in mean water level close to the shore to which the waves are travelling, and in other conditions there is a fall.

The investigation of the motion and equilibrium which is the basis of physics can if applied to the ocean achieve rapid increase in understanding of physical processes that are the basis of great engineering projects. The annual bill for delay and damage to ships through waves is tremendous. It has already been shown that great saving can be made by using wave predictions up to 3 or 4 days ahead to plan optimum routes for vessels crossing the North Atlantic Ocean. An experiment with 1000 ships of the U.S. Military Sea Transport Service crossing the North Atlantic Ocean showed an average saving of 10% of the passage time. It would be very difficult for naval architects and engineers to save 10% as cheaply, but there is no doubt that better knowledge of ship motion in relation to waves will help in the design of hulls for particular speeds in specified wave conditions. The millions spent in protecting coasts and dredging harbours, in retaining beach material in front of the eroding land that needs its protection, and in removing it from the places where it will hinder navigation, will surely be spent more effectively when we have closer understanding of the circulation of water on the beaches and at the bottom of the estuaries. The potential gains might at first glance seem small in comparison with the weeks and even months that were saved by sailing ships using optimum routes, but are in fact greater under the greater intensity of trade and competition to-day.

There is also much to be gained from studying longer changes in sea level, ranging from the 1- to 2-minute oscillations that produce troublesome range action in harbours, through the 'tidal waves' caused by seismic disturbances of the sea floor or intense meteorological disturbances of the surface, to the tides themselves produced by the variations at points in the earth's surface of the attractive forces of the moon and sun. The 1- to 2-minute oscillations probably get their energy from 'surf beats', due to the varying accelerations of water as groups of high and low waves break on gently sloping beaches. Whether or not these are really the cause of the harbour oscillations is not yet known, though the question is of such importance that a recent Navigational Congress urged that research be undertaken to establish the cause, to predict severe attacks and to devise alleviating and remedial measures. Such work might be judged to be hydraulic engineering rather than oceanography but both sciences will lose if they do not overlap, oceanography going up the rivers to high-water mark and hydraulics down the estuaries to low-water mark.

Submarine earthquakes usually affect areas with dimensions large compared with the depth of water and most of the energy goes into waves that are long compared with the depth of water. They travel at the speed $\sqrt{(gh)}$ where g is the acceleration due to gravity and h the depth of water, and in the deep ocean this generally amounts to between 400 and 500 miles per hour. This is, of course, the speed of the wave not of the water particles. The four or five waves travelling across the deep ocean are likely to be only 1 to 2 feet high with as much as 100 miles between the crests,

and quite imperceptible to ships in deep water. It is only when they move into shallow water that the decrease in speed produces destructive amplification. Seismic surges travel right across the ocean—those from the Chilean earthquake had a destructive effect on the coast of Japan. Their rate of travel and time of arrival can be predicted with great accuracy with the help of information from seismographs and islands near the source, but the size of the wave and the violence with which it will run on to particular parts of distant coasts depends very much on the direction of approach, offshore bottom topography, protecting reefs and local resonances. Much more work is needed to prepare the way for full predictions.

Coastal floods produced by atmospheric disturbances are just as serious. A disturbance moving over the sea tends to produce a wave and if the disturbance travels at the speed of a free wave $\sqrt{(gh)}$ there is a near-resonance effect in which the wave, receiving energy all the time, grows very large. Winds and pressure disturbances do not travel fast enough to excite such resonance over the deep ocean, but they can do it over coastal shelves and shallow seas. The disastrous hurricane surges on the Atlantic coast of U.S.A. are caused by hurricanes travelling at just the right speed over a sufficient length of shallow water. There have been very disastrous waves on the east coast of India, caused by disturbances moving at the critical speed over the shallow water at the head of the Bay of Bengal. The same mechanism is responsible for smaller rises in water level on our own west coast, and even in the English Channel a change in pressure or wind, almost too small to be noticeable but travelling at the right speed, produces a wave in which the water rises and falls several feet in as many minutes, enough to cause considerable confusion on a holiday beach or among boats moored in an estuary. The serious floods of January 1953 on our east coast were due to a surge that travelled down the east coast and round the southern part of the North Sea to the coasts of Holland and Germany. The height of the surge seems to have been due mainly to the piling up of water before the strong north wind. Everywhere in the ocean the water level is sensitive to barometric pressure. In the deep ocean and along a straight coast the sea behaves just like an inverted barometer and there is almost a one-to-one relationship between the rise in barometer and depression of the sea, or the fall in barometer and rise in the sea. In narrower seas and estuaries the response is somewhat less because of the inertia of the water. At London Bridge a rise in air pressure corresponding to 1 inch on the mercury barometer causes a depression of 7 inches in the height of high water.

On a longer scale there are changes in the monthly and annual means of sea level obtained by smoothing out the tides over these periods. Round our own coasts the monthly mean sea level is some 15 to 20 cm higher in autumn than in spring. Most of it can be accounted for by the difference in temperature and pressure, but there is also some evidence that there is more water in the ocean at the end of summer than at the end of winter, when more is stored in ice, snow, lakes and rivers on land. The seasonal change is generally much smaller in the tropics, but at the head of the Bay of Bengal there is a difference of as much as 165 cm between the two monsoon seasons. The fluctuations in mean annual level are generally less than ± 5 cm from year to year, but may be as much as 10 cm. Except in high latitudes they show an upward trend, the water rising relative to the land. In the south of England the rate of rise corresponds to about 6 inches in 100 years. This begins to be large enough to be important to engineers since so much of our population and industry lives on ground below the level of the highest spring tides. Nothing can be done to stop surges of water due to storms or seismic disturbance, but measures can be taken to lessen the damage when accurate predictions are available. It might be possible to devise economical schemes to deflect some of the energy away from particularly vulnerable and important places. At least it is important to be able to form the best possible estimates of what might happen over a range and coincidence of weather and coastline. Not all parts of the world are sinking relative to the sea. We do not know how much water is being added to the ocean as a whole through the melting of ice caps and glaciers, or whether the temperature of the ocean as a whole is rising enough to cause appreciable expansion, but as far as we can see its level as a whole is rising at a rate of about 4 inches in 100 years. Subsidence and uplift of the land itself plays a considerable part. Tide gauges in nothern countries such as Finland, Canada and Alaska show a rise of land relative to the sea, and there is no doubt that this is due to continuing uplift of land which is recovering from the great load of ice which it carried during the last glacial period. The northern parts of Finland are rising at a rate of 3 feet in 100 years and on the shores of Hudson's Bay the rise is twice as fast.

Sea level changes from place to place as well as time to time. The main factors are differences of temperature, density, atmospheric pressure, wind, currents and waves. The well-known difference between the two ends of the Panama Canal may be largely due to differences of density. In the rainy season which affects the Gulf of Panama mean sea level is 30 cm higher at the Pacific end than in the more saline Atlantic water. In the dry season the difference of level is very small. In the Straits of Dover a tidal current of 3 knots towards the North Sea would make the sea level at Calais 50 cm higher than at Dover because of the sideways thrust of the earth's rotation. The tilt reverses with the tidal stream, but there is enough residual non-tidal current towards the North Sea to leave a difference of a few centimetres.

No body of water, however small, nor the earth itself, is left undisturbed. An American writer has said that geodesists who have hopefully used 'sea level' as a reference, and oceanographers who equally hopefully have used 'fixed bench-marks', have finally bumped into one another going backwards. Tides have been measured in the Taylor Model Basin, which is about $\frac{1}{2}$ mile long. They were about three-thousandths of an inch and would have been nearer four-thousandths if the earth had been solid instead of flexible. Earth tides are known to have an appreciable effect on ocean tides, and the water tides flowing in to coastal areas tilt the neighbouring land to a small but measurable extent. Interference of surface patterns of ocean waves causes fluctuations in mean pressure on the sea bed large enough to produce measurable microseismic oscillations in the middle of the continent. They can be used to obtain information about the ocean waves.

The greatest interplay of force and movement in the oceans is in the currents and general circulation of water, and the processes, including the interchange of energy between the ocean and atmosphere, are complex. The winds and water movements influence the density layering of the ocean, altering its thermodynamic potential as well as its response to the winds. The ocean currents driven by the winds and density differences do much to determine where most heat is fed to the atmosphere and so help to set the pattern of the winds. It is the effect of the earth's rotation and the way it increases with latitude that makes the currents stronger on

the western side of each ocean. Although the subject has been of great interest to sailors, travellers and scientists for hundreds of years, we have only skimmed the surface. We have good charts of the average current at the surface but do not understand the eddies and day-to-day changes found in them. We have only just discovered that the water movements at great depths can be almost as fast and variable as those at the surface and have no theory to explain it. We must learn much more about how these fluctuating movements are generated before we can understand and predict changes that affect navigation as well as the slower changes that affect climate or the conditions on fishing grounds.

It seems almost incredible in such days of unparalleled scientific enterprise, discovery and achievement that the advantages of a more detailed understanding of a great sector of our environment, much more important to mankind than Space, are overlooked. The problems themselves are difficult, but challenging and exciting, and the knowledge to be gained has almost unlimited applications and yet the work is neglected. The book *Modern Physics for the Engineer*, published by McGraw-Hill in 1954, devotes 36 pages to astrophysics and 24 to the earth below the sea, but not a word about the sea itself. The discoveries and developments of the past few years are striking enough to secure some mention next time, but the omission is indicative and representative of a general lack of appreciation of the potential value of what eminent scientists who never go out of a laboratory call a 'fringe subject'. The oceans are still a bit more than a fringe, and the trade and defence of our country will not be maintained without scientific investigation of the ocean a bit more advanced than the accurate observations and plottings of the last century.

New Books

Digital computers and nuclear reactor calculations. By W. C. SANGREN. (New York and London: John Wiley, 1960.) Pp. xi + 208. 68*s*.

More digital computer time has been devoted to nuclear engineering than to any other branch of science; most of this time has been spent solving parabolic and elliptic partial differential equations, and simultaneous ordinary differential equations. The present book provides a brief account of the phenomena which the various equations describe, enumerates the salient features of automatic computers and the principles of their proper use, and indicates the form in which the equations are presented to a computer for solution.

The first chapter describes the problems facing nuclear design engineers, and there follow chapters on Digital computers, Programming, and Numerical analysis. The discussion is restricted, necessarily, but it is designed to make the reader aware of what computers can and cannot do; for example, the desirability of an error analysis for every computation is emphasized. The value of automatic coding facilities is stressed and the potentialities of ALGOL are indicated. The vital statistics of American machines are summarized (though the figures appear rather conservative now), and users of English machines will be interested to learn that the routine manufacture of reliable magnetic tape has been achieved abroad. Surprisingly, the importance of a facility for accumulating exact scalar products of vectors is not mentioned, and the use of direct methods for solving elliptic difference equations is discouraged.

In Chapter 5, equations for reactor poisoning are presented, and a programme for their solution is described in detail. Chapter 6 is devoted to diffusion and age-diffusion equations: these are approximations to the fundamental Boltzmann transport equation for neutrons, and give rise to the criticality eigenvalue problem and the associated inhomogeneous boundary value problem. Methods of attacking the transport equation itself are the subject of Chapter 7, most space being given to a Monte Carlo technique. The final chapter is a miscellany, and touches on reactor kinetics, burnout, shielding, and heat conduction. From the viewpoint of the general reader, these chapters would be improved by a diagram illustrating the disposition of fuel-cans, moderator, etc., and by a short account of the physical processes which are exploited in a reactor. For example, no comment is made on the role of thermal neutrons, or on the meaning of individual terms in the transport equation.

The attitude to computing manifest in this book is exemplary, and such work could reasonably become prescribed reading for diploma courses on numerical analysis. (Companion volumes on fluid dynamics and on guided missiles would be useful.) The following general remarks deserve publicity: 'the value of any program is greatly enhanced by a good write-up', p. 12; 'it is unlikely that one-of-a-kind machines will again play a major part in computing', p. 24; 'economics seems to dictate . . . the increased use of large central computer facilities', p, 27; 'debugging at the console wastes a great deal of machine time', p. 45; 'there will always exist a class of problem, for any computer, where it is economically sensible (for the operator at the console) to make a decision based upon an intermediate result', p. 47.

In view of the book's virtues, the poor grammar and flabby style of writing are to be deplored. Inadequate proof reading has left a residue of trivial errors, which occur more frequently than one per fifteen pages.

D. W. MARTIN

Calculation of the behaviour of rubber-covered pressure rollers

by G. J. PARISH, B.Sc., A.Inst.P., British Cotton Industry Research Association, Shirley Institute, Manchester

MS. received 30*th December* 1960, *in revised form* 14*th March* 1961

Abstract

Data are given which enable the properties of nips between metal and rubber-covered pressure rollers to be calculated from the parameters of the system. The properties with which the calculation is primarily concerned are the nip width and the mean pressure in the nip, but the principle can be extended to include the peak pressure and the distribution of pressure through the nip. The calculation is based on the well-known Hertzian formula and on empirical relations, which express the important effects of the rubber-cover thickness.

The data refer not only to roller systems in which the loading is uniform, but also to non-uniform nips, in particular to those which show what is probably the commonest cause of non-uniformity, roller deflection.

Although the calculations refer primarily to nips between one hard roller and one relatively soft, covered, roller, they are directly applicable to nips between two identical covered rollers and may be applied, within limits, to systems in which the rollers are dissimilar in properties.

1. Introduction

RUBBER-COVERED pressure rollers are widely used industrially for the processing of material in sheet form. The material may be squeezed either between two rubber-covered rollers or between a rubber roller and one of a hard material, such as metal, ebonite or granite. These materials are sufficiently harder than rubber for it to be safe to assume that they retain their cylindrical shape in the nip. In either case, then, the properties of the nip depend on the properties of the rubber-covered roller or rollers alone.

Strictly speaking, this will only be true if there is no compressible material between the rollers. However, if the indentation of the rubber cover in the nip is appreciably greater than the change in thickness of the processed material, the presence of this material can have very little effect on the nip properties. This condition is certainly satisfied by the great majority of textile fabrics and must also be satisfied by many other materials. In these circumstances it is reasonable to ignore the material which is being processed and to assume that the nip properties are determined by the roller alone.

Ideally, a full knowledge of the way in which the pressure is distributed through the nip is required in order to define the nip properties. However, the symmetry and general similarity shown by the results of experiments on a variety of nips (Parish 1958) suggest that the mean pressure alone should define the nip properties with sufficient precision for practical purposes. This has been found to be the case in the mangling of textiles (Moss and Parish 1959); a knowledge of the mean pressure should be equally useful in other processes, although it might be desirable also to know the maximum pressure. This paper deals with the calculation of these pressures from given roller parameters, considering first the contact of a rubber-covered roller with a hard roller and subsequently the contact of two rubber-covered rollers.

2. Theory of contact between a hard and an elastic roller

This theory has been discussed previously (Parish 1958). For a homogeneous elastic roller of rubber, Poisson's ratio being taken as $\frac{1}{2}$, and assuming plane strain and uniform loading, the nip width is given by

$$h_0^2 = \frac{3}{2} \cdot \frac{WD}{\pi E} \tag{1}$$

where $2h_0$ is the nip width, W is the applied load per unit length of the rollers, E is the Young's modulus of the rubber, and

$$\frac{1}{D} = \frac{1}{D_1} + \frac{1}{D_2} \tag{2}$$

D_1 and D_2 being the roller diameters.

The pressure distribution through the nip is

$$P(x) = P_m(1 - x^2/h_0^2)^{1/2} \tag{3}$$

where x represents distance measured from the centre line of the nip ($-h_0 \leqslant x \leqslant h_0$), $P(x)$ is the pressure at point x, and P_m is the maximum value of this pressure ($x = 0$).

The mean pressure is given by

$$\bar{P} = W/2h_0 \tag{4}$$

$$= (\pi/4)P_m \tag{5}$$

from Eqn (3).

Eqns (1)–(5) enable the properties of the nip to be calculated from the parameters W, E, D_1 and D_2. The Young's modulus E can be obtained from measurements with an indentation hardness tester, the relationship having recently been discussed by Gent (1958). Ideally, the indentation should be measured with a dead-load instrument on a rubber slab, but reasonable results can be obtained with a 'pocket' instrument, even on the curved surface of a roller, if it is used with care.

The above conditions apply only to a homogeneous rubber roller. If the roller consists of a relatively thin cover on a hard core, Eqns (1) and (3) no longer hold (Parish 1958). An additional parameter, introducing the cover thickness b, is involved. The nip width $2h$ is smaller than the value $2h_0$, calculated on the assumption that the roller is homogeneous, and an experimental curve relating the nip width ratio h/h_0 to the parameter $2h/b$ is shown in Fig. 3 of the paper by

Parish (1958). A second experimental curve, in Fig. 2 of that paper, expresses the failure of Eqn (3) by a plot of $P_m/\bar{P}$ against the same parameter. For purposes of computation the first of these relations is better expressed by a graph of h/h_0 against h_0/b, and this is shown in the present Fig. 1. This graph enables a calculation procedure based on

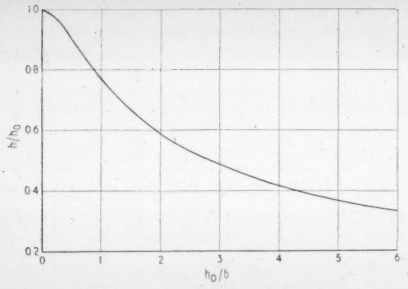


Fig. 1. Variation of h/h_0 with h_0/b.
This figure is based on Fig. 3 of a previous paper (Parish 1958).

Eqn (1) to be extended to a roller with a cover of thickness b. Most rubber-covered rollers are constructed with a layer of ebonite between the metal shell and the outer rubber layer. The ebonite is sufficiently hard to be regarded as part of the shell, and the cover thickness b is the thickness of the soft rubber layer alone.

3. Extension to conditions of uneven load distribution

A procedure for calculating the nipping properties of rollers would be of limited value if it could only be applied to those systems in which the distribution of load is uniform along the length of the rollers. Many industrial machines have some degree of non-uniformity of load distribution, the commonest cause of which is roller deflection. Deflection arises from the fact that the inward thrust is applied to the bearings of the rollers, whereas the outward thrust is distributed over the roller working face. The result is that the nip loading, and hence the nip pressure, vary in an unknown manner from point to point along the length of the rollers. The length of the roller face will be denoted by $2l$ and the distance along the face, measured from the centre, by y (then $-l \leqslant y \leqslant l$). Eqn (1) can no longer be used directly, since the value of $W(y)$ is not known, but only the average value $\bar{W}$. The important point is that, when roller deflection is present, the variation in separation of the hard parts of the rollers along their length must be balanced by a variation in the indentation of the rubber cover or covers. The extent of the variation in load distribution will depend on the rubber properties, but in all cases the variation in indentation must be equal to the combined deflection of both rollers. This deflection, in the plane of the roller axes, will be denoted by $z(y)$, relative to the centre of the face (so that $z(0) = 0$). If the rollers, in addition to showing deflection, have some surface camber (crown), then the residual deflection is given, of course, by the difference between the true deflection and the camber.

The extension of the calculation to deflecting rollers therefore necessitates the determination of rubber indentations. Recently, Loo (1958) has derived theoretically the indentation of a homogeneous roller, but the treatment has not yet been applied to a covered roller. It has been necessary, therefore, to rely entirely on experiment.

The rubber indentation is measured, in the plane containing the roller axes, by measuring the difference in separation of the hard parts of the two rollers when the surfaces are brought just into contact without load, and when load is applied. The experiment must, of course, be made on a roller system in which deflection is negligible. Measurements have been made in this way of the indentation of several rubber covers when loaded, rotating, against a steel roller in the apparatus on which the pressure distribution measurements (Parish 1958) were carried out. In each case the measured indentation δ was compared with a 'geometrical' indentation d given by

$$d = h^2/D \tag{6}$$

and the ratios δ/d were found to lie close to a single curve when plotted against the parameter h_0/b. For purposes of computation a more convenient plot is of the function $\delta D/b^2$ against h_0/b, and this is shown in Fig. 2.

With this additional relation it is possible to extend the calculation to non-uniformly loaded systems. The information required for the calculation now includes a knowledge of the deflection of each roller. The resistance to bending of a roller is provided by the metal parts alone and the deflection may be calculated by the usual methods, as is described, for example, by Gough (1947) and by Sturken and Verner (1955). Calculation, however, does not always give reliable results. Where possible, measurement is to be preferred, and a simple instrument, the Shirley bowl deflection indicator, has been developed for this purpose. A description of this instrument and a brief discussion of roller deflection are included in a paper by Moss (1959).

Returning to the calculation of nip properties, the essential problem is to determine the way in which the load $W(y)$ must vary along the rollers in order to satisfy the roller deflection conditions. The necessary relationships are contained in Eqn (1) and Figs 1 and 2, but a solution can only be reached by a procedure of repeated trials. The author has used nomograms to facilitate the calculation.

4. Calculation procedures

4.1. *Hard roller/rubber-covered roller: uniform loading*

Under these circumstances the calculation is straightforward. The parameters required are W, D_1, D_2, rubber hardness E, and b. The value of h_0 is calculated and Fig. 1 is used to obtain h. This gives the mean pressure directly; if the value of P_m is required, it can be determined from Fig. 2 in the paper by Parish (1958).

4.2. *Hard roller/rubber-covered roller: non-uniform load*

The required parameters are as given above (except that the load per unit length is now the average value $\bar{W}$), with the addition of the combined roller deflection $z(y)$. The calculation procedure is to assume a value for W_c (i.e. $W(y)$ for $y = 0$), and to use this value to compute δ_c using Eqn (1) and Fig. 2. Since the roller deflections are known, this gives $\delta(y)$ for any value of y. Hence $W(y)$ may be determined for a suitable number of points and the average value compared with the given value of $\bar{W}$. The correct value of W_c is that which makes these two averages equal.

In most cases it is probably unnecessary to compute $W(y)$ for more than one point other than the centre, and a slightly different procedure may be more convenient. The shape of

roller deflection may be represented with reasonable accuracy by

$$z(y) = z_m(y/l)^2, \tag{7}$$

z_m being the deflection at the ends of the roller face.

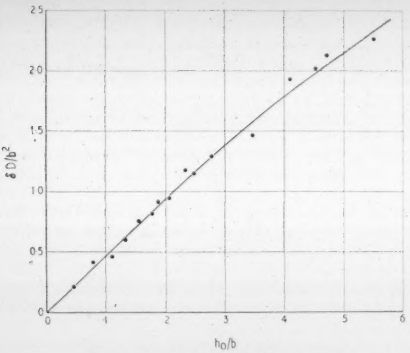


Fig. 2. Variation of $\delta D/b^2$ with h_0/b.
The points are experimental values.

It is convenient to choose the second point as that for which $z(y) = z_m/2$, that is, for $y = \pm l/\sqrt{2}$. Now, it is not unreasonable to represent the variation in load distribution also by a second-order equation

$$W(y) = W_c + K(y/l)^2. \tag{8}$$

(This equation clearly cannot be strictly correct, but it is sufficiently accurate for most practical purposes.)

Hence

$$3\bar{W} = W_c + 2W_s, \tag{9}$$

where the subscript s denotes the side position, for $y = \pm l/\sqrt{2}$.

If this equation is assumed to hold, the appropriate value of W_s is fixed once the value of W_c is chosen, and δ_s may be calculated from W_s in the same way that δ_c is calculated from W_c. The check on the correctness of the value assumed for W_c then lies in the difference between δ_s and δ_c.

4.3. *Identical rubber-covered rollers*

By identical is meant that the rollers are of the same diameter and have covers of equal hardness and thickness. From symmetry it is clear that, with uniform loading, in the nip the surface of each roller must lie on a plane normal to the plane containing the roller axes. Each roller then behaves as though it were pressed into contact with a rigid plane. The calculation proceeds in a manner exactly similar to that outlined in Section 4.1, except that the diameter D is now set equal to the diameter of one rubber roller ($D_2 = \infty$).

Virtually the same conditions apply with non-uniform loading. The surface of contact may now be curved in the plane of the roller axes, but will still be flat in the direction normal to this plane. Each small element along the length of each roller thus behaves again as though in contact with a rigid plane. The calculation proceeds as in Section 4.2, again with D set equal to the diameter of one roller alone, but now, since the roller covers are identical, they must share equally the variations in indentation required by the combined deflections of the two rollers. That is, the difference $\delta(y) - \delta_c$ of indentation of each roller must now be half of the combined deflection $z(y)$. (Note that the deflections of the two rollers need not necessarily be equal, since the properties that govern roller rigidity are not involved in the identity conditions required here.)

The procedure of this subsection applies exactly only if the rollers are identical. It can, of course, be applied to give a less accurate solution if the rollers do not differ too greatly; under these conditions the average values of the parameters would be used.

4.4. *Dissimilar rubber-covered rollers*

In general the rollers may differ in diameter, in hardness and in cover thickness. Subscripts 1 and 2 will be used to denote the two rollers.

4.4.1. *Uniform loading.* Although the nip width $2h$ must have a unique value (for a given load), the theoretical nip widths, $2h_{01}$, $2h_{02}$, will not be equal if $b_1 \neq b_2$. The surface of contact will no longer be plane; it may, without undue simplification, be assumed cylindrical and of diameter D_3. It is then necessary to find the value of D_3 to fit the following conditions:

$$\left.\begin{aligned} h_{01}^2 &= \frac{3}{2}\frac{WD'}{\pi E_1} \\ h_{02}^2 &= \frac{3}{2}\frac{WD''}{\pi E_2} \end{aligned}\right\} \tag{10}$$

where

$$\left.\begin{aligned} \frac{1}{D'} &= \frac{1}{D_1} + \frac{1}{D_3} \\ \frac{1}{D''} &= \frac{1}{D_2} - \frac{1}{D_3} \end{aligned}\right\} \tag{11}$$

and the values h_{01}, h_{02} are so related to b_1, b_2 that $h_1 = h_2$. The procedure then must be one of trial and error, made by choosing values of D_3 until these conditions are satisfied.

In the special case where $b_1 = b_2$ the problem is much simpler, since then $h_{01} = h_{02}$ and the general equation for contact of two elastic rollers may be used, both Poisson's ratios being set equal to $\frac{1}{2}$,

$$h_0^2 = \frac{3WD}{2\pi}\left(\frac{1}{E_1} + \frac{1}{E_2}\right). \tag{12}$$

(It may be noted that Eqns (10) and (11) also lead to Eqn (12) for $h_{01} = h_{02}$.)

4.4.2. *Non-uniform loading.* The general problem becomes very complicated. Although the total indentation at any value of y is known from the roller deflections, for a chosen value of $(\delta_1 + \delta_2)_c$, the relative proportions of $\delta_1(y)$ and $\delta_2(y)$ are not known; neither is it safe to assume that the value D_{3c} at the centre of the roller faces applies for all values of y.

The only feasible procedure is to assume Eqn (8) to hold, so that W_s is fixed when W_c is chosen. D_{3c} and D_{3s} must then be chosen so that the necessary conditions are satisfied at each point, as described above. It still remains to determine that the difference $(\delta_1 + \delta_2)_s - (\delta_1 + \delta_2)_c$ satisfies the deflection conditions. If not, another value of W_c must be chosen and the calculation repeated.

In the special case where $b_1 = b_2$ the problem is simpler. Eqn (12) can then be used for h_0, and the value of

$D\{(1/E_1) + (1/E_2)\}$ is the same for all values of y; the calculation is only a little more complicated than the similar calculation with one hard and one covered roller in contact.

4.5. *Varying rubber thickness*

The calculation procedures described above have all assumed a constant rubber thickness along the length of the roller or rollers (ignoring the minute differences which result if roller surface camber is present). A substantial variation in rubber thickness can be used to counteract the effects of roller deflection, and a calculation can be made to determine what thickness variation is required for a particular application.

The usual methods of correcting for roller deflection (cambering of the roller surface, skewing of the roller axes, etc.) all have for their object the attainment of a uniform distribution of *load* along the length of the nip. This would appear to be an unnecessarily stringent condition; the load may be allowed to vary provided that the nip width is made to vary in the same manner, so that the mean pressure is kept constant. This result can often be obtained by varying the thickness of the rubber cover along the length of the roller. A cover which is thicker at the ends of the roller face than in the centre is required, and this can be obtained, while maintaining a cylindrical outer surface to the roller, by cambering the roller mandrel (Parish 1955). This method of correcting for roller deflection has several advantages over the conventional surface camber.

It would be possible, in theory, by suitably shaping the mandrel, to obtain a constant mean pressure at every point along the length of the roller, but the procedure would be very complicated. For practical purposes it is sufficient to apply a parabolic camber profile (in many instances even an approximation to this parabola will suffice), and to assume that the load is parabolically distributed (Eqns (8) and (9)).

The calculation is again made by repeated trials, first assuming a value for W_c and calculating $\bar{P}_c$ and δ_c as indicated previously. Since W_s is fixed by the chosen value of W_c and Eqn (9), the value of h_{0s} can be determined, and since $\bar{P}_s$ and $\bar{P}_c$ are to be made equal, $2h_s$ is also fixed. The value of b_s is thus determined by means of Fig. 1, and δ_s may be found. The difference $\delta_s - \delta_c$ is compared with the combined roller deflection z_s to provide the check on the correctness of the value assumed for W_c.

The difference in cover thickness $b_s - b_c$ indicates the amount of camber to be put on the mandrel. The behaviour of the roller at other loads and other cover thicknesses, or for other positions along the face can now be calculated by the method described in Section 4.2 above, using the appropriate values of cover thickness.

5. Conclusions

The calculation procedures described here have been applied over the past few years to a wide variety of textile mangles, and results in good agreement with measurable parameters have been obtained. One such measurable parameter is, of course, the effect of the nip on the material being processed, but often this will provide only a qualitative estimate of the behaviour of the system. The quantity which provides the best comparison with calculation is the nip width. The calculations are based on empirical relationships derived from a system in which the rollers were rotating, and the calculated results therefore refer to such 'dynamic' conditions. Measurements of nip width can be made under similar conditions by a photographic method (Moss 1959), and a modification of this method can be applied with the rollers stationary, as can the well-known carbon paper and similar methods. Such measurements show that there is very little change in nip width with speed, but there is a somewhat greater difference between measurements made with the rollers rotating and with them stationary. 'Static' nip widths are larger than 'dynamic' ones and the difference generally lies between 5 and 10%. In comparing calculated nip widths with static measurements, allowance must be made for this difference. With this allowance made, calculated and measured values have been found to be in good agreement, generally to within 5%.

It must be remembered, however, that the calculations are based on empirical relations which show significant departures from a theoretical treatment. It was concluded elsewhere (Parish 1958) that these differences were due to the non-linear elasticity of the rubbers employed. If this is so, the successful application of the results to many other rollers implies that all the covers possessed the same order of non-linearity. This would seem to be quite possible with the relatively limited range of rubber hardness involved, roughly 55° to 90° B.S. This range will include most requirements for pressure roller covering, but the calculation should be applied with caution to hardnesses outside this range and to other covering materials.

Acknowledgment

The author wishes to acknowledge the permission of the Director of the Shirley Institute to publish this paper.

References

GENT, A. N., 1958, *Trans I.R.I.*, **34**, 46.

GOUGH, J. B., 1947, *Paper Trade J. (TAPPI Sect.)*, **125**, 213.

LOO, T-T., 1958, *J. Appl. Mech.*, **25**, 122 (see also discussion by LUBKIN, J. L., 1958, *J. Appl. Mech.*, **25**, 638).

MOSS, E., 1959, *J. Text. Inst.*, **50**, 654.

MOSS. E., and PARISH, G. J., 1959, *Textile Engineering Processes*, Ed. A. H. NISSAN (London: Butterworths), p. 248.

PARISH, G. J., 1955, British Patent 795, 523.

—— 1958, *Brit. J. Appl. Phys.*, **9**, 158.

STURKEN, R. C., and VERNER, D. A., 1955, *Tappi*, **38**, 676.

Photosensitivity and speed of response in cadmium sulphide

by D. SHAW, B.Sc., D.I.C., Ph.D., A.Inst.P., Research Laboratory, Associated Electrical Industries Ltd., Rugby

MS. received 21*st March* 1961

Abstract

The photoconductivity and photocurrent decay time at 875 *foot candles illumination have been measured for over ninety crystals of cadmium sulphide of widely different photosensitivities. The same parameters were also measured over the illumination range* 0·007 *to* 875 *foot candles for a smaller number of crystals. At* 875 *foot candles a clear correlation between photoconductivity and decay time was observed. The performance of all the crystals was limited by trapping. Evidence for two types of trap was obtained.*

Introduction

THE practical interest in the photoconductive properties of cadmium sulphide stems from the discovery by Frerichs (1947) of the particularly high photosensitivity to visible radiation found in suitably prepared cadmium sulphide crystals. It is generally recognized (Rose 1956) that the photocurrent for a given excitation rate is controlled by the lifetime of the excited photocarriers, which is defined as the time spent by the carriers in the conduction and/or valence bands between generation and recombination. Thus if uniform optical generation in a photoconductor creates f electron hole pairs in unit volume in unit time, then steady state conditions exist when the recombination rate of free carriers is equal to their generation rate. This condition is expressed by the equation

$$f = n/\tau_n = p/\tau_p \qquad (1)$$

where n and p are the free electron and free hole concentrations respectively and τ_n and τ_p are the respective mean lifetimes. Eqn (1) assumes that the thermal equilibrium carrier concentrations are negligible compared with n and p, which is a condition usually fulfilled in large band gap photoconductors.

In practical applications of a photoconductor two important parameters governing its performance are the photosensitivity and speed of response. On the basis of a photoconductor containing only recombination centres within the forbidden band and one predominant carrier, the photocurrent would show an exponential rise and decay with a time constant equal to the free carrier lifetime. Thus from Eqn (1) a high photocurrent at a given rate of excitation also entails a large response time and vice versa. However, as trapping effects must be considered in the case of transients, such as photocurrent decay, the results of this simple model need modification to allow for trapping effects. Rose (1951, 1956) has shown that while trapping has no effect on the steady state, thus Eqn (1) remains valid, the decay time constant is lengthened by a factor n_t/n or p_t/p, where n_t and p_t are respectively the number of electrons and holes held in traps, for steady state free carrier concentrations n and p, and which are released before the free carrier concentration falls to 1/e of its steady state value. The rise time is similarly lengthened due to the filling up of the empty traps.

Despite the considerable literature on the photoconductive effect in cadmium sulphide, there is very little information concerning response time of the photocurrent and in particular its relation to the rate of excitation and photosensitivity. Bube (1956b) investigated ten crystals using a constant rate of excitation and concluded that the response time was independent of photosensitivity except for the most sensitive crystals. The work described in this paper is an extension of Bube's (1956b) earlier measurements through the investigation of a much larger number of crystals. In addition the effect of varying the rate of excitation is also studied.

Calculation of photoconductor performance

Equation (1) enables a relationship between the photocurrent, illumination and lifetime to be derived. We shall consider specifically a single crystal of undoped, homogeneous cadmium sulphide and assume a negligible hole conductivity (Rose 1957). Let F be the total number of photons absorbed per cm^2 of illuminated surface. Each photon is assumed to create one electron hole pair. Then in the steady state

$$F = \int n(z)\frac{dz}{\tau(z)} \qquad (2)$$

where z is the depth measured normal to the illuminated surface. $n(z)$ and $\tau(z)$ are the free electron concentration and mean lifetime respectively at a depth z and the integral extends over the thickness of the crystal. If the crystal is a rectangular slab of width a, length b and thickness c, the observed photocurrent under uniform illumination is

$$i = q\mu aV\int_0^c n(z)\frac{dz}{b} \qquad (3)$$

for a voltage V applied across the crystal length by means of ohmic end contacts. μ is the electron conductivity mobility and q the electronic charge.

In order to combine Eqns (2) and (3) it is necessary to know or assume the dependence of $n(z)$ and $\tau(z)$ upon z. Making the assumption, which will be justified later, that $\tau(z) = \tau$, a constant, permits $n(z)$ to be eliminated between Eqns (2) and (3) to yield

$$\sigma/F\tau = \mu q \qquad (4)$$

where $\sigma = (i \times b)/(a \times V)$, which will be defined as the photoconductivity of the crystal. Physically it is the electrical conductivity per square of illuminated surface: it is a characteristic solely of the crystal and not of the surface or electrode geometry.

Kroger *et al.* (1955) obtained a value of 210 $cm^2 v^{-1} s^{-1}$ for the Hall electron mobility at 300° C due to lattice scattering in cadmium sulphide. They showed that this value held for both surface and volume conduction. Later work by Miyazawa *et al.* (1959) and by Portis (1959) has given excellent confirmation of this value for the Hall mobility. In general the Hall and conductivity mobilities are unequal but, for lack of information, equality will be assumed in the present case.

If trapping is present then τ is no longer equal to the observed photocurrent decay time. The latter exceeds τ by a factor determined by the amount of trapping. In practice it proves more convenient to measure the time t taken to decay to half the steady state photocurrent. Allowing for this change from τ to t and substituting values for μ and q into Eqn (4) gives

$$\sigma/Ft \leqslant 4{\cdot}8 \times 10^{-17} \text{ mho cm}^2. \qquad (5)$$

The term σ/Ft will be defined as the performance of the crystal. This is basically the same definition used by Rose and Lampert (1959). Evidently maximum performance, corresponding to no trapping, is given by equality in Eqn (5). It is also logical to define σ/F as the photosensitivity of the crystal.

The evaluation of F is straightforward in the case of monochromatic illumination through the use of Dutton's (1958) results on optical absorption and reflection in cadmium sulphide. For white light illumination this data can be used to show that a pure crystal completely absorbs all wavelengths shorter than a particular one and is essentially transparent to longer wavelengths. The precise cut-off wavelength is about 0·52 μ and depends upon the crystal thickness. F is thus set equal to the number of photons entering the crystal with wavelengths less than the cut-off.

Crystal preparation

The crystals for this investigation were grown by the sublimation of phosphor grade cadmium sulphide powder within a quartz tube. A carrier gas, composed of nitrogen and hydrogen sulphide in the proportions 10 : 1, conveyed the cadmium sulphide vapour to the crystal growth zone. The sublimation temperature of the powder charge was 1080° C and the crystals grew in a temperature zone centred about 850° C.

The colour of the crystals was pale yellow and their habit ranged from thin plates and laths to long slender needles. Dark electrical conductivities were in the range 10^{-10} to 10^{-8} mho cm^{-1} with photoconductivities of less than 10^{-6} mho under 20 f.c. illumination.

Mass spectrographic analysis was used to check the purity of the grown crystals. Potassium and chromium were the only two elements detected. Important impurities such as the group 1B and 3B metals were below the detection limit of the instrument. The amounts and detection limits for the various impurity elements are summarized in Table 1.

Table 1

Element	K	Ag	Cu	Al	Ga	Cr
Concentration atoms $cm^{-3} \times 10^{-16}$	7–70	<0·5	<0·9	<4	<0·8	0·28–2·8

The presence of chlorine was masked by the instrument blank but its concentration was less than 8×10^{17} atoms cm^{-3} and probably less than 8×10^{16} cm^{-3}.

In order to achieve as wide a variation in crystal photosensitivity as possible, sample batches from the grown crystals were subsequently fired in different atmospheres. Ten batches were selected, each batch containing of the order of 100 crystals. The firing time was thirty minutes for all batches. Table 2 summarizes the firing conditions.

Table 2

Batch		1	2	3	4	5	6	7	8	9	10
Firing temp. ° C		700	700	700	700	800	800	800	800	475	500
Gas ambient	% H_2S	—	5	30	100	—	5	30	100	—	—
	% N_2	100	95	70	—	100	95	70	—	—	—
	% Air	—	—	—	—	—	—	—	—	100	100

Dark conductivities for batches 1 to 8 ranged from approximately 0·1 mho cm^{-1} to less than 10^{-11} mho cm^{-1}. It is reasonable to attribute this wide spread of values to variations from stoichiometry. An interesting feature of the crystals from these batches was the relatively abrupt rise in photosensitivity with decreasing dark conductivity. Thus for dark conductivities less than about 2×10^{-3} mho cm^{-1} the ratio of photocurrent to dark current was at least a factor ten at 875 f.c. illumination compared to a ratio of less than two for dark conductivities greater than 2×10^{-3} mho cm^{-1}.

The dark conductivities for the air fired batches showed little change from their original values. The principal effect of the air firing was to increase the photosensitivity.

Measurements

Approximately ten crystals were selected from each batch for measurement. Pale yellow crystals only were chosen in order to ensure negligible impurity photoconductivity. There were, in fact, very few crystals in any of the batches showing colouration due to impurity absorption. Ohmic contacts were made to the individual crystals by soldering with indium. Typical crystal dimensions were:
$0{\cdot}05 < a < 0{\cdot}2$ cm, $0{\cdot}2 < b < 0{\cdot}5$ cm, $0{\cdot}005 < c < 0{\cdot}02$ cm.

Values of σ and t were measured for all the selected crystals which had a light to dark current ratio of at least ten at the maximum test illumination of 875 f.c. These parameters were also measured over the illumination range 0·007 to 875 f.c. for a smaller number of crystals. The source of illumination was a calibrated Ediswan d.c. Pointolite lamp operating at a candle power of 94·3 cd and a colour temperature of 2900° K; t was measured using a chopper disk which gave an on–off time ratio of 14·2 : 1 with two chops per revolution. The decay was displayed and measured on an E.M.I. type WM8 oscilloscope. In general the rise and decay times were very similar for a given crystal; neither the rise or decay of photocurrent could be represented by a simple exponential. All measurements were made at room temperature in room air. The crystal characteristics proved quite stable under these conditions. The measured values of σ and t have an estimated accuracy to $\pm 10\%$.

Dutton's (1958) absorption data lead to a wavelength cut-off of about 0·515 μ for $c = 0{\cdot}01$ cm. Doubling the thickness raises the cut-off to 0·52 μ. In view of the present range in crystal thickness F has been calculated for a cut-off wavelength of 0·515 μ. Thus 875 f.c. illumination at a colour temperature of 2900° K yields $F = 1{\cdot}40 \times 10^{15}$ photons $cm^{-2} s^{-1}$. Losses by reflection of the incident radiation at the crystal surface were allowed for by assuming a constant value of 0·2 for the spectral reflection coefficient (Dutton 1958). An extension of the cut-off wavelength to 0·52 μ raises F to $1{\cdot}52 \times 10^{15}$ photons $cm^{-2} s^{-1}$. Thus F is not critically dependent on c. Indeed the absorption data (Dutton 1958) show that all wavelengths less than 0·51 μ

will be completely absorbed within a depth of 10 μ. This means that 80–90% of the absorbed photons are absorbed within a 10 μ thick surface layer. Clearly the processes occurring within this layer will largely decide the photoconductive properties. Because of its very small thickness it appears reasonable to assume substantially uniform properties within this layer. This is the basis for setting $\tau(z) = \tau$ in deriving Eqn (4). Thus τ is taken to be the mean lifetime of free electrons within this thin surface layer. Substituting $F = 1{\cdot}40 \times 10^{15}$ photons cm^{-2} s^{-1} into Eqn (5) gives

$$\sigma/t \leqslant 6{\cdot}8 \times 10^{-2} \text{ mho s}^{-1}. \qquad (6)$$

Results

875 *f.c. illumination*

Figure 1 records σ plotted against t at 875 f.c. for all the selected crystals, 92 in number. For clarity no attempt is made to distinguish between crystals from different batches. Generally, however, batches 4 and 8 gave crystals with low σ and t and batches 1 and 5 crystals with high values for these two parameters. The diagonal line corresponds to maximum performance as given by equality in Eqn (6).

The data of Fig. 1 show two distinct features. The first is the diffuse but definite trend for σ to increase with t over

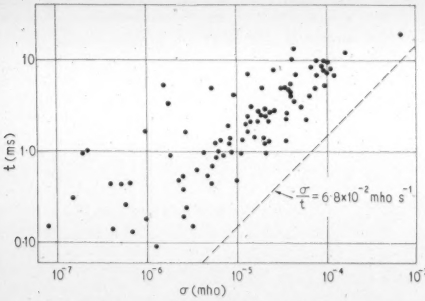


Fig. 1. Variation of photoconductivity, σ, with decay time t at 875 f.c.

the range $0{\cdot}10 < t < 10$ ms. This differs from Bube's (1956b) conclusion that σ was largely independent of t in the range $0{\cdot}30 < t < 4$ ms. The distribution shown in Fig. 1 suggests that the diffuseness becomes more marked as σ decreases. There is, however, no indication of a lower limit to t. The second point is that all the crystals lie to the left of the maximum performance line and thus conform to Eqn (6). It can be seen that the observed performances are mainly between one fifth and one tenth of the maximum value.

Illumination range 0·007 *to* 875 *f.c.*

A total of thirteen crystals were studied in this illumination range. Seven crystals were from batch 1 and two each from batches 4, 8 and 10. In Fig. 2 the observed performance σ/Ft is plotted against σ for three of these crystals. The dependence of σ on F for the same three crystals is shown in Fig. 3. It will be seen from this last figure that, at low F, σ shows a super-linear dependence. Eleven of the crystals exhibited this property with a power law dependence on F ranging from 1·05 to 1·67. Of the two remaining crystals one (batch 10) was linear while the other (batch 1) was sub-linear at low F.

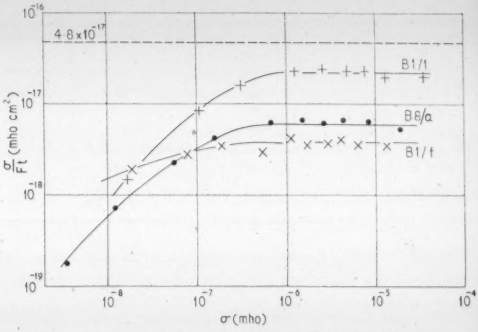


Fig. 2. Variation of performance σ/Ft with photoconductivity.

All the super-linear crystals gave performance characteristics similar in form to those in Fig. 2, namely a constant level of performance until σ fell to between 10^{-6} and 10^{-7} mho, below which the performance steadily decreased. In all cases the constant performance level was less than the maximum given by Eqn (5). It will also be seen from Fig. 3

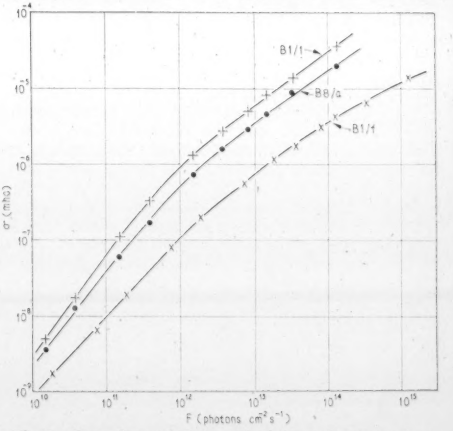


Fig. 3. Dependence of photoconductivity on the photon flux F. The curves for B1/1 and B8/a have been displaced horizontally to the left by a factor 10.

that as F decreases, the change from a sub-linear to a super-linear variation occurs when $10^{-7} < \sigma < 10^{-6}$ mho. This characteristic was shown by all eleven super-linear crystals.

In contrast the performances of the two remaining non-super-linear crystals decreased steadily over the whole range of σ. The performance curve in both cases, however, did show a marked change in gradient at $\sigma \sim 10^{-7}$ mho. A final point of interest concerns the lower curve in Fig. 3 which, while super-linear at low F, shows a square root dependence at high F.

Discussion

The performance data of Figs 1 and 2 are in agreement with Eqn (5) with the highest observed performances being about 50% of the calculated maximum. It is possible to account for differences of this order in terms of a reduction in electron mobility due to impurity scattering. The analysis of Conwell and Weisskopf (1950), Erginsoy (1950) and Conwell (1952) shows that an impurity concentration of 10^{18} atoms cm^{-3} will give in cadmium sulphide a combined lattice and impurity mobility of approximately one-half the lattice mobility. Reference to Table 1 shows that foreign impurities could be present at this concentration level but that this would be an upper limit. Anion and cation vacancies will also give rise to impurity scattering but an application of the method of Kroger *et al.* (1953) to the firing conditions for batches 1 to 8 gives a maximum vacancy concentration of 10^{17} cm^{-3}. Vacancy scattering should therefore be comparatively unimportant. The difference between the calculated and observed performance maxima can thus be attributed to trapping, to the influence of impurities on the mobility or to a combination of these two effects. Trapping, however, must be responsible for greater differences.

The general dependence of the performance on σ shown in Fig. 2 exhibits two distinct trapping effects. For $\sigma > 10^{-6}$ mho a constant but smaller than maximum performance implies that the ratio n_t/n is constant. Below $\sigma = 10^{-7}$ mho the continuous deterioration in performance corresponds to n_t/n rising as σ falls due to decreasing illumination. n_t is decreasing but at a lesser rate than n. The transition between these two trapping modes, due presumably to two distinct trapping levels, occurs when the steady state electron Fermi level (Rose 1957) lies between 0·36 and 0·42 ev below the conduction band, using a thickness of 10 μ for the conducting layer.

The change from a sub-linear to a super-linear variation of σ with F also occurs as the steady state electron Fermi level traverses the same energy range. The present results are, however, insufficient to decide on the significance of this parallel effect. Bube (1956a) has found that in 'pure' cadmium sulphide crystals, super-linear behaviour only occurred when the steady state electron Fermi level lay between 0·6 and 0·8 ev below the conduction band. Super-linearity was only observed therefore at temperatures of 150° C or above. Room temperature super-linearity has, however, been found in CdS : Cu (Bube 1955), CdS : Cu : Cl (Wright 1958) and CdS : Ag : Cl (Avinor 1959). In all three cases the impurity or doping concentrations were in the range 10^{18} to 10^{19} atoms cm^{-3} and were large enough to make the crystals red sensitive. The location of the steady state electron Fermi level in these instances is unknown.

In the present work the spectral response of the photoconductivity was typical of that associated with 'pure' cadmium sulphide crystals and the only significant impurity was potassium at a concentration of the order of 5×10^{17} atoms cm^{-3}. Further work is clearly desirable to establish the conditions under which super-linearity occurs.

As the illumination increases the photoconductivity will eventually be determined by direct recombination between free electrons and free holes. This situation will arise when the free carrier concentrations exceed the concentration of recombination centres. Recombination will clearly be bimolecular so that $\sigma \propto F^{1/2}$. If, as is likely, the recombination process is radiative, the free carrier lifetime, $\tau_n = \tau_p = \tau$, will be quite independent of the defect structure and can be calculated by the method of van Roosebroeck and Shockley (1954). Applying this method to the measured optical absorption data (Dutton 1958) yields a value of $4{\cdot}7 \times 10^{-9}$ cm^{-3} s^{-1} for the radiative recombination probability in cadmium sulphide at 300° K. The refractive index was assumed constant and equal to 2·6. The ratio of the effective mass to the electron rest mass, for both electrons and holes, was assumed to be 0·27 (Kroger 1955). This probability gives a capture cross section of $4{\cdot}7 \times 10^{-16}$ cm^2 for the recombination of a free hole with a free electron. With present impurity concentrations in cadmium sulphide radiative recombination is unlikely until the free carrier concentration equals or exceeds 10^{17} cm^{-3}. Taking this figure as the average concentration in the 10 μ thick surface layer together with the radiative capture cross section enables the photon flux, required to produce this carrier concentration, to be calculated from Eqn (5). Thus $F \sim 10^{22}$ photons cm^{-2} s^{-1} and is clearly outside the optical excitation range. It is also of interest to calculate the density of recombination centres so that direct recombination is still dominant for $F = 1{\cdot}4 \times 10^{15}$ photons cm^{-2} s^{-1}. The result is that the density must be of the order of 10^{12} cm^{-3} or less. These calculations show that, with present impurity concentrations, the photoconductivity is unlikely to be determined by direct carrier recombination. In particular the dependence $\sigma \propto F^{1/2}$, shown by the lower curve in Fig. 3, cannot be attributed to direct recombination.

Conclusions

The main result arising from this investigation is that the performance up to 875 f.c. illumination in cadmium sulphide crystals is always less than the theoretical optimum value. Observed performances generally lie between a factor five to ten below the theoretical value. This loss in performance has been interpreted as due to trapping effects which are revealed by an increase in response times. From the practical viewpoint this loss is quite significant. A gain in response time by a factor five to ten, without loss of photosensitivity, would enable cadmium sulphide photocells to be used in many applications from which they are at present barred. In principle, trapping could be eliminated at least over the working range of the crystal. There is, however, the possibility that the trapping levels may be associated with the recombination centres which determine the photosensitivity. Therefore any attempt to improve performance should first be directed to identifying the recombination and trapping centres.

Acknowledgments

The author is indebted to Mr. J. T. Anderson for the preparation and heat treatment of the crystals investigated. Thanks are also due to Mr. M. Wilson for the mass spectrographic analyses.

References

AVINOR, M., 1959, *Philips Res. Rep.*, **14**, 211.

BUBE, R. H., 1955, *Proc. Inst. Radio Engrs, N.Y.*, **43**, 1836.

—— 1956a, *Photoconductivity Conference* (New York: John Wiley), p. 575.

—— 1956b, *J. Appl. Phys.*, **27**, 1237.

CONWELL, E. M., 1952, *Proc. Inst. Radio Engrs, N.Y.*, **40**, 1331.

CONWELL, E. M., and WEISSKOPF, V. F., 1950, *Phys. Rev.*, **77**, 388.

DUTTON, D., 1958, *Phys. Rev.*, **112**, 785.

ERGINSOY, C., 1950, *Phys. Rev.*, **79**, 1013.
FRERICHS, R., 1947, *Phys. Rev.*, **72**, 594.
KROGER, F. A., VINK, H. J., and VAN DEN BOOMGAARD, J., 1953, *Z. Phys. Chem.*, **203**, 1.
KROGER, F. A., VINK, H. J., and VOLGER, J., 1955, *Philips Res. Rep.*, **10**, 1.
MIYAZAWA, H., MAEDA, H., and TOMISHIMA, H., 1959, *J. Phys. Soc. Japan*, **14**, 41.
PORTIS, A. M., 1959, *J. Phys. Chem. Solids*, **8**, 326.
VAN ROOSEBROECK, W., and SHOCKLEY, W., 1954, *Phys. Rev.*, **94**, 1558.
ROSE, A., 1951, *R.C.A. Review*, **12**, 362.
—— 1956, *Photoconductivity Conference* (New York: John Wiley), p. 3.
—— 1957, *Progress in Semiconductors*, Vol. 2, (London: Heywood), p. 111.
ROSE, A., and LAMPERT, M. A., 1959, *Phys. Rev.*, **113**, 1227.
WRIGHT, D. A., 1958, *Brit. J. Appl. Phys.*, **9**, 205.

New Books

Histological techniques for electron microscopy. By D. C. PEASE. (New York and London: Academic Press, 1960.) Pp. xii + 274. Price $7.50.

Although a satisfactory method of cutting sections of tissue thin enough for electron microscopy was worked out quite early in the history of the subject, no detailed treatment of the art has been available until now. The initial 'proof' by von Ardenne in 1939 that it was mechanically impossible to cut sections much thinner than 1 micron was followed by ill-advised attempts to cut at extremely high speeds. In 1948 Pease and Baker in the United States and Bretschneider in Holland showed that the type of microtome used for optical sections could readily be modified so as to cut electron-transparent sections in an entirely orthodox way. Since then a great deal of experimentation has gone on into the best methods of preparing and cutting different biological materials.

One of the originators of the technique has now set down full details of the various procedures in this very clear and well arranged text. He starts with the statement, born of long experience, that 'to make successful electron micrographs of tissues one must preserve the specimens with skill, section them with art, use an electron microscope with understanding, and do photography with facility'. The nine chapters of his book are planned to purvey all the information required to equip the microscopist in these various arts. They deal with the collection, fixation and embedding of tissues, the sectioning and mounting of sections, before going on to the operation of the electron microscope and the taking and processing of the photographs. References to the literature are given at the end of each chapter, and some notes on the available bibliographies and on the published proceedings of conferences are given in an appendix. A second appendix gives a useful list of suppliers of equipment and materials.

The treatment is clear and, where the reviewer could test it, comprehensive. It is illustrated not only with typical micrographs of various tissues, but also by some remarkably clear and informative photographs of the main procedures of handling, mounting and sectioning specimens. With this guide in his hands, anyone new to the subject should be able to learn with a minimum of error how to use an electron microscope to good advantage in biological and medical research. 'When in trouble, the novice is urged to return to the pertinent parts of the text to search for overlooked details', is the author's good advice. He also stresses that the biologist trained in the preparative methods of optical microscopy must get used to the refinement of method and of manual dexterity which electron microscopical work demands. The regular cutting and mounting of sections 200 Å thick and only $\frac{1}{10}$ mm^2 in area is an art as much as a science. Dr. Pease's handbook is likely to become the standard *vade-mecum* for all who wish to master it.

V. E. COSSLETT

The influence of an electric field on the growth of copper whiskers

by T. HOFMAN*, M.Sc., J. MAZUR†, D.Phil., F.Inst.P., J. NIKLIBORC*, D.Phil., and J. RAFAŁOWICZ†, M.Sc.,
* Physics Department, the University, Wroclaw, Poland. † Low Temperature Laboratory, Institute of Physics, Polish Academy of Sciences, Wroclaw, Poland

MS. received 19*th January* 1961, *in revised form* 6*th April* 1961

Abstract

The influence of an electric field on the growth of copper whiskers obtained by Brenner's method, has been studied. There is a marked directional tendency during the growth process and it is thought that ions may play a part in the phenomenon.

1. Introduction

THE problem of the influence of an electric field on the growth of copper whiskers has to the best of our knowledge not been studied.

There are few references concerning other metals. Arnold (1956) tried to determine the effect of electric and magnetic fields on the rate of growth of tin whiskers on tin plated specimens and no effect has been found. Gomer (1958), and Melmed and Gomer (1959) have found an exponential variation with electric field of the rate of growth of mercury whiskers. Lewowski and Sujak (1960) have carried out investigations on the effect of the electric field on the growth of molybdenum oxide crystals.

In order to fill the gap we have decided to study the influence of the electric field on copper whiskers obtained by Brenner's method.

As a first step we started our observations on CuI for the following reasons: (i) it is comparatively simple to obtain copper whiskers from CuI, (ii) CuI has been widely studied by Brenner (1959) and Morelock and Sears (1959).

2. Experimental arrangement

The experiments have been carried out by using the standard Brenner method consisting of a copper boat and upper copper electrode; a high tension supply generated the necessary electric field. The apparatus is shown schematically in Fig. 1.

The distance between the electrodes was 20 mm and the applied electric field varied from several hundreds to about 2000 volts; higher fields were deliberately avoided, because they caused glow discharge which immediately destroyed the growing whiskers (Fig. 2).

During all our experiments the only variable parameter was the electric field; the other controlled parameters—temperature, the reaction time, the quantity of CuI, the geometry of the apparatus—were maintained constant.

The observations were carried out both without electric field and with electric field; the boat and the upper electrode were alternately positively and negatively charged.

3. Experimental observations

3.1. *Experiments without electric field*

In the absence of the electric field whiskers were never found on the upper electrode; whiskers are formed in the lower boat and do not show any directional tendency: the angles between the whiskers and the surface of the bottom of the boat are random (Fig. 3).

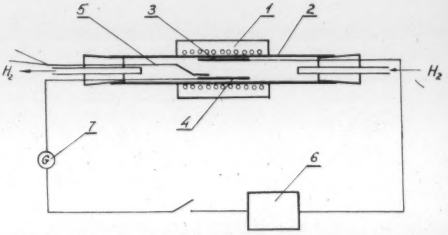


Fig. 1. 1, tube furnace; 2, quartz tube; 3, upper electrode; 4, boat filled with CuI; 5, thermocouple; 6, high tension supply; 7, galvanometer.

3.2. *Experiments with electric field*

When the upper electrode was positively or negatively charged whiskers were always formed on it and this fact seems to indicate that the material for the whisker structure is present in the form of ions or polarized molecules.

The characteristic feature of the whiskers obtained on the upper electrode is their perpendicular direction, i.e. at right angles to the electrode surface (Fig. 4). From this photograph it is clear that the whiskers follow the lines of force of the electric field.

When the lower boat, filled with CuI, was positively or negatively charged, the whiskers were always growing; the striking majority of the whiskers shows a distinct directional effect (Fig. 5). The whiskers grown in the charged boat are much shorter than those grown without electric field; it seems that the electric field shortens the length of the whiskers. This assumption is confirmed by the fact that the tips of the whiskers grown in the presence of the electric field are in general 'unravelled' or blunt ended (Fig. 6), whereas the tips of the whiskers grown without electric field are mostly thinner and taper; even when the whiskers are thicker, their tips are sharp (Fig. 7). Some whiskers produced probably at the beginning of the electric discharge terminate in a ball

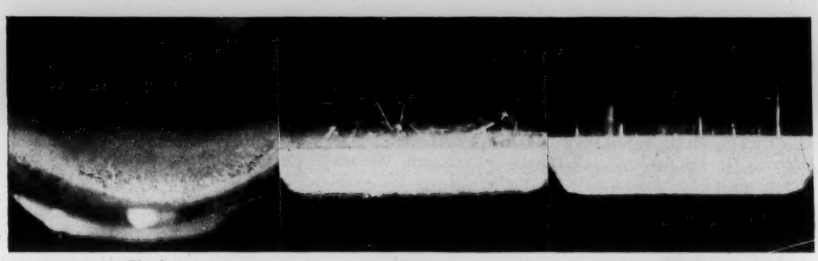


Fig. 2. Fig. 3. Fig. 4.

Fig. 5. Fig. 6. Fig. 7.

Fig. 8. Fig. 9(*a*). Fig. 9(*b*).

(Fig. 8). The ends of some whiskers are forked (Figs 9(*a*), (*b*)). These branched whiskers occur whether or not an electric field is present but occur more frequently when it is applied.

4. The measurement of the electric current

When the boat was empty, the hydrogen was flowing and the temperature was exactly the same as that during the process of whisker growth, the value of the current was of some microamperes, whereas when the boat was filled with CuI and whiskers were growing, the other conditions being the same as before, the current was of the order of some milliamperes. This observation supports the view previously expressed that during the growth of the whiskers ions play a very important part.

5. Conclusions

The problem of the effect of the field direction still remains open. It is necessary to continue the present research by using electron microscope and x-ray technique in order to study the exact profile of the whiskers, curtailed by the electric field, and the microstructure of the branching.

Further experiments are in progress and the results will be published elsewhere.

References

ARNOLD, S. M., 1956, *Techn. Proc. 43rd Ann. Convention Amer. Electroplaters Soc.*, Bell Telephone System Monograph 2635.

BRENNER, S. S., 1959, *Acta Met.*, **7**, 519.

COLEMAN, R. V., and SEARS, G. W., 1957, *Acta Met.*, **5**, 131.

GOMER, R., 1958, *J. Chem. Phys.*, **3**, 457.

LEWOWSKI, T., and SUJAK, B., 1960, *Z. Naturf.*, **15**, 89.

MELMED, A. J., and GOMER, R., 1959, *J. Chem. Phys.*, **30**, 587.

MORELOCK, C. R., and SEARS, G. W., 1959, *J. Chem. Phys.*, **4**, 926.

Magnetic method for the estimation of ferrite in stainless steel welds

by G. C. CURTIS, Ph.D., A.Inst.P., and J. SHERWIN, B.Sc., U.K. Atomic Energy Authority (Development and Engineering Group) Research and Development Branch, Windscale, Cumberland

MS. received 30*th January* 1961, *in revised form* 10*th March* 1961

Abstract

A method is described of determining, with an estimated probable error of $\pm\frac{1}{2}$, *the percentage of ferrite in a stainless steel weld by measurement of the saturation intensity of magnetization of the weld. A calibration curve for* 18/8/1 *stainless steel is given, and is correlated with curves obtained with iron in wax specimens. There seems no reason why the range of* 2 *to* 18% *ferrite content actually measured should not be extended, using modified apparatus.*

Introduction

In the welding of some stainless steels, duplex structures of ferrite and austenite are formed and it may be important to have a means of determining the percentage of ferrite present. Several methods are available: in the variant of the point-count method used at Windscale, a suitably treated portion of the specimen is examined microscopically and the proportions estimated by the interception of ferrite grains on a random line superimposed on the field of view.

This method is, however, inherently slow and tedious. The possibility was therefore examined of utilizing the fact that ferrite is ferromagnetic and austenite paramagnetic.

The most relevant literature discovered (Simpkinson and Lavigne 1949, De Barr 1953, Fleischmann 1954) included no description of a suitable experimental method. It was at first hoped that a simple inductance bridge could be used, and that the increase of inductance obtained by inserting a specimen into one of the arms would be a measure of its ferrite content. To test the method, artificial specimens were prepared by suspending in wax two iron powders—M.C.P. (Mond Carbonyl Powder) and S.I.P. (Swedish Iron Powder)—of widely different particle size, but there was no correlation between the results obtained from these two samples. It was then decided to attempt to develop a method independent of factors like particle size, and this note describes the work done.

Theory of method using saturation intensity of magnetization

As the saturation intensity of magnetization depends only upon the nature and amount of ferromagnetic material present and not upon factors such as particle size or shape, the intensity of magnetization of a weld specimen should give a measure of the amount of ferrite present.

Let a field H be applied along the axis of the specimen, the flux being B. Then at saturation

$$B - H = 4\pi I_s = cf \qquad (1)$$

where c depends only on the nature of the ferromagnetic substance and f is its fraction by volume in the specimen.

By suddenly withdrawing a search-coil connected to a ballistic galvanometer, in presence and in absence of the sample, we can measure B and H separately. However, we are obtaining cf from the difference of two much larger quantities, thus magnifying any error. If we obtain deflections by suddenly annulling H (switching off an electromagnet) the same objection applies. In the latter case the main pulse could be backed off electrically, but difficulties arising from differences in waveform could be expected. If, on the other hand, we hold the search-coil fixed and suddenly remove the sample we obtain a deflection θ proportional to $B - H$ directly. Thus

$$f = k\theta \qquad (2)$$

where k is a constant depending on c and on the apparatus.

Apparatus

The biggest item of apparatus, an electromagnet, was borrowed from a 'Magnaflux' crack detector: the ease with which this instrument can be converted to a ferrometer is noteworthy, because many small metallurgical laboratories already have this type of equipment available. The highest and most uniform field, in the gap, could not be used without drilling a specimen withdrawal hole through one pole-piece, at the risk of cold-working it. The specimen and coil were therefore placed alongside the gap, which was plane-parallel and fairly wide. The usable field was thus 2200 oersted, and was found to give a sufficiently close approach to saturation to permit reasonable accuracy (Fig. 1).

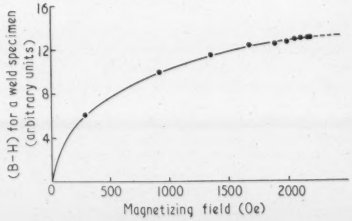


Fig. 1. Showing approach to saturation.

The coil had 1050 turns wound on a glass tube $1\frac{1}{2}$ in. long and just over $\frac{1}{4}$ in. in diameter; it was connected to a ballistic galvanometer. The weld specimen was 3 in. long and $\frac{1}{8}$ in. in diameter, which was the maximum diameter obtainable without including the base plate of the weld. The specimen was withdrawn sharply from the coil and the deflection was noted.

The field was measured with an accurately made coil in conjunction with a mutual inductor.

Results

Several M.C.P.- and S.I.P.-in-wax specimens were tested, as well as several 18/8/1 stainless steel weld specimens. The iron content of the wax was obtained by weighings and the

ferrite content of the steel by the point count method with an accuracy to $\pm\frac{1}{2}\%$. All the specimens were of the same length.

Figure 2 is a plot of ferrometer reading (deflection per unit cross section of specimen) against percentage of iron by

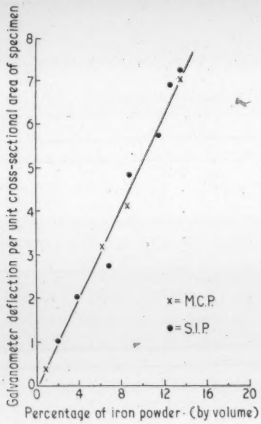


Fig. 2. Calibration of ferrometer with iron-in-wax specimens.

volume for the M.C.P. and S.I.P. powders in wax. It verifies Eqn (1) and confirms that the reading is independent of particle size.

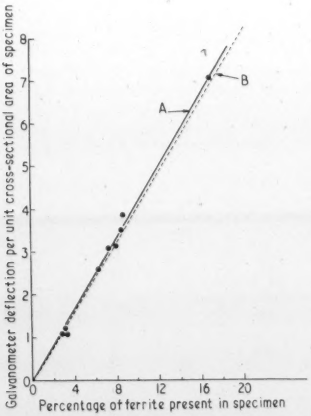


Fig. 3. Calibration of ferrometer for 18/8/1 steels.
A, actual weld (determined by point-count method).
B, iron-in-wax (ordinate multiplied by composition factor 0·78).

In Fig. 3, curve A shows the relation between ferrometer reading and point count estimate for the 18/8/1 weld specimens.

Discussion

The slopes for the iron-in-wax specimens (Fig. 2) and weld specimens (Fig. 3, curve A) do not agree. This can be explained as follows.

According to Seitz (1940), chromium and nickel atoms when substituting in the iron lattice act as ferromagnetic atoms, the chromium contributing 0·22, the nickel 0·6 and the iron 2·2 effective electrons per atom. Let W_1, W_2 and W_3 be the percentages by weight of chromium, nickel, and iron in the ferrite, and A_1, A_2 and A_3 the respective atomic weights. Let d be the density of the ferrite and L Avogadro's number. Then the chromium contributes $0{\cdot}22\,LdW_1/A_1$ effective electrons per hundred cubic centimetres of ferrite. The total saturation induction will therefore be

$$\left(0{\cdot}22\frac{W_1}{A_1}+0{\cdot}6\frac{W_2}{A_2}+2{\cdot}2\frac{W_3}{A_3}\right)Ld \quad \text{in similar units.}$$

The densities of different ferrites will agree closely, so the factor Ld may be ignored in comparing readings. Assuming that the ferrite approximates to the nominal composition of the welding electrode $W_1 = 18\%$, $W_2 = 8\%$, $W_3 = 74\%$, we obtain 3·07 units for the expression in brackets. For pure iron, $W_1 = W_2 = 0\%$ and $W_3 = 100\%$, giving 3·94 units.

Assuming also that the M.C.P. and S.I.P. powders are pure iron, we should expect to have to multiply every ordinate of Fig. 2 by 3·07/3·94 or 0·78 to make it applicable to 18/8/1 weld specimens. This gives a line which has been plotted as B in Fig. 3, agreeing well with line A drawn through the experimental points determined by the point-count method. Even if it were assumed that the weld included a proportion of material derived from the 18/13/1 baseplate the agreement would be little affected since the factor of 0·78 deduced above must be much nearer the truth than the factor of 0·74 derived from the baseplate composition $W_1 = 18\%$, $W_2 = 13\%$, $W_3 = 69\%$.

The departure of the points from the straight line in Fig. 3 may be partly due to errors of the point-count method. Thus an estimate of $\pm 0{\cdot}5$ for the probable error in the determination of the percentage of ferrite by the ferrometer method may be on the high side.

It appears then that the ferrite content of welds of different stainless steels could be found if their compositions were known.

Conclusions

The method described should give a rapid and reliable measurement of the ferrite content of a stainless steel weld. It has the advantage of giving directly the mean value throughout a specimen of convenient size, thus eliminating the need to make a number of determinations at different points and then to average them.

Acknowledgments

Thanks are due to Professor E. C. Stoner for his interest and advice in the early stages of the work, and to Sir William Cook, Managing Director, and Dr. H. Kronberger, Director of Research and Development of the Development and Engineering Group, United Kingdom Atomic Energy Authority, for permission to publish.

References

De Barr, A. E., 1953, *Soft Magnetic Materials used in Industry* (The Institute of Physics Monographs for Students).

Fleischmann, W. F., 1954, *Weld. J., Easton, Pa*, **33**, 4335.

Seitz, F., 1940, *Modern Theory of Solids*, Chap. VIII (New York: McGraw–Hill).

Simpkinson, T. V., and Lavigne, M. J., 1949, *Metal Progr.*, **55**, 164.

The matching of Pierce guns to tunnels

by C. J. MILNER, M.A., Ph.D.(Cantab.), F.Inst.P., and K. J. AUSBURN, M.Sc.(Lond.), D.I.C., A.Inst.P.,
School of Physics, The University of New South Wales, Australia

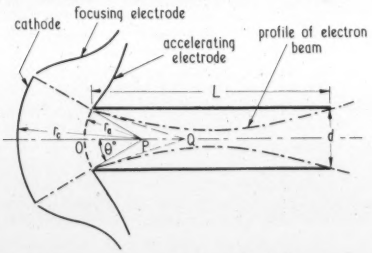
MS. received 25th November 1960, in revised form 6th March 1961

Abstract

All Pierce guns which shoot the maximum electron current through two identical apertures collinear with the axis of the gun have a unique ratio of cathode to anode radius. The value of this unique ratio is given for four important cases. A simple design procedure and relevant numerical data are given.

Introduction

In a recent concise review article Meltzer (1959) refers briefly to the design of Pierce electron guns for such devices as klystrons and mentions the effect of the aperture lens on gun design. The gun design problem for the axially symmetric case is illustrated in the Figure. One



Typical Pierce gun and tunnel. Conical beam requires OQ = $L/2$. Strip beam requires OQ = $L/4$.

wishes to pass the maximum possible current, at a given accelerating potential, through the two circular apertures of diameter d and spacing L. In a klystron the apertures are in the wall of a resonator and the region between them is assumed field free for the gun design problem. However, allowance must be made not only for the diverging action of the anode aperture mentioned by Meltzer, but also for the beam spread due to space charge repulsion (Watson 1927) occurring in the field free region. A design procedure, together with relevant families of curves, has been described by Spangenberg (1948). A greatly simplified design procedure results from recognition of the fact that all Pierce guns which shoot maximum current through a tunnel at a given potential have, to a paraxial approximation, a unique value for the ratio of cathode radius to anode radius.

Optimum values of gun parameters

This result for cylindrical tunnels has been published by Kleen (1958) who obtained graphically a unique value for the optimum cathode to anode radius of $r_c/r_a = 2{\cdot}15$. The present authors have independently derived this result (solving the final equation by numerical methods) and have obtained the value $r_c/r_a = 2{\cdot}14$. The corresponding value of the Langmuir–Blodgett (1924) function α^2 is 0·946 compared with Kleen's value of 0·98.

Similar results have been derived by the present authors for the following cases not previously discussed.

(*a*) Cylindrical tunnel with the aperture lens effect removed by the presence of a fine grid over the anode aperture. Here $r_c/r_a = 1{\cdot}47$ and $\alpha^2 = 0{\cdot}1891$.

(*b*) Strip beam matched to a pair of identical slits, without a grid over the anode aperture: $r_c/r_a = 2{\cdot}85$, $\beta^2 = 2{\cdot}59$ where β^2 is the corresponding Langmuir–Blodgett (1923) function.

(*c*) Strip beam with a fine grid over the anode slit: $r_c/r_a = 1{\cdot}71$, $\beta^2 = 4/9$.

Design procedure

If, for example, a current I is to be passed through a tunnel of length L and diameter d the procedure is as follows.

The necessary accelerating potential V is calculated from

$$I = (1{\cdot}084)^2 2\pi\varepsilon_0 \left(\frac{2e}{m}\right)^{1/2} V^{3/2} \left(\frac{d}{L}\right)^2 \qquad (1)$$

(Pierce 1949) where e and m are the electronic charge and mass respectively and ε_0 is the permittivity of free space. Rationalized M.K.S. units are used throughout.

If, following Kleen, the current (1) is equated to the expression for current from a Pierce gun ($1/4\pi$ times the spherical-diode current (Langmuir and Blodgett 1924) multiplied by the solid angle of the segment), then, to the paraxial approximation (see Figure)

$$\sin \tfrac{1}{2}\theta = \frac{d}{4r_a},$$

one obtains

$$L^2 = 18(1{\cdot}084)^2 \alpha^2 r_a^2 \,. \qquad (2)$$

Substitution of the prescribed optimum value for α^2 allows the required value of r_a to be calculated from (2). From the prescribed value of r_c/r_a, the cathode–anode spacing follows directly:

$$r_c - r_a = \left(\frac{r_c}{r_a} - 1\right) \frac{L}{4 \cdot 60(\alpha^2)^{1/2}} \,. \qquad (3)$$

Moreover the cathode diameter d_c is simply

$$d_c/d = r_c/r_a \,. \tag{4}$$

These expressions apply whether or not a grid is used, provided the appropriate values of α^2 and r_c/r_a are used.

For the case of a beam of current I (per unit length of slit) between two slits, each of width d, in planes L apart corresponding equations to (1), (2) and (3) are

$$I = 16\varepsilon_0\left(\frac{2e}{m}\right)^{1/2} V^{3/2}\left(\frac{d}{L^2}\right) \tag{1'}$$

$$L^2 = 36\beta^2 r_a{}^2 \tag{2'}$$

$$r_c - r_a = \left(\frac{r_c}{r_a} - 1\right)\frac{L}{6(\beta^2)^{1/2}} \,. \tag{3'}$$

Since there is a unique value of cathode radius to anode radius the electrolytic trough analogue for the determination of the electrode shapes for the Pierce gun may then be of a universal character. The anode and cathode radii may be fixed and only θ, the semi-angle of the beam, need be altered. The potential distribution along the beam boundary will be the same in all cases and may be computed once for all. Tabulated values are given in the appendix.

Appendix

It is convenient to tabulate here computed values of the potential distribution that must be established along the beam edge in the four cases quoted, namely the conical-beam and strip-beam systems, with and without a grid covering the anode aperture. In a spherical diode or sector thereof, i.e. in a conical beam, the potential relative to cathode at radius r is proportional to $(\alpha^2)^{2/3}$ while in a cylindrical diode or sector thereof, i.e. in a strip beam, it is proportional to

$$\frac{(\beta^2)^{2/3}}{(r_c/r)^{2/3}}$$

where α^2 and β^2 are the relevant functions of r_c/r as computed by Langmuir and Blodgett. In the appended Tables, for each of various round-number values of r_c/r, there is given (i) the value of $(r_c - r)/(r_c - r_a)$, that is, the distance from the cathode as a fraction of the anode spacing, and (ii) the value of the potential V at the point as a fraction of the anode potential V_a.

Computed potential distribution in Pierce guns matched to tunnels as described in the text, and consequently having values of r_c/r_a and α^2 or β^2 (at the anode) as stated

Conical, with no grid

$\frac{r_c}{r_a} = 2{\cdot}14$ $(\alpha^2)_a = 0{\cdot}946$

$\frac{r_c}{r}$	$\frac{r_c - r}{r_c - r_a}$	$\frac{V}{V_a}$
1·05	0·089	0·0186
1·10	0·171	0·0469
1·15	0·245	0·0797
1·20	0·313	0·116
1·25	0·375	0·154
1·30	0·433	0·194
1·35	0·486	0·236
1·40	0·536	0·279
1·45	0·582	0·323
1·50	0·625	0·369
1·60	0·704	0·462
1·70	0·772	0·558
1·80	0·834	0·655
1·90	0·889	0·755
2·00	0·938	0·857
2·10	0·983	0·959

Strip, with no grid

$\frac{r_c}{r_a} = 2{\cdot}85$ $(\beta^2)_a = 2{\cdot}59$

$\frac{r_c}{r}$	$\frac{r_c - r}{r_c - r_a}$	$\frac{V}{V_a}$
1·1	0·140	0·0458
1·2	0·257	0·108
1·3	0·356	0·173
1·4	0·440	0·239
1·5	0·514	0·304
1·6	0·578	0·367
1·8	0·685	0·488
2·0	0·771	0·600
2·2	0·841	0·705
2·4	0·899	0·803
2·6	0·949	0·895
2·8	0·991	0·981

Conical, with grid

$\frac{r_c}{r_a} = 1{\cdot}47$ $(\alpha_2)_a = 0{\cdot}1891$

$\frac{r_c}{r}$	$\frac{r_c - r}{r_c - r_a}$	$\frac{V}{V_a}$
1·05	0·149	0·0544
1·10	0·284	0·137
1·15	0·408	0·233
1·20	0·521	0·338
1·25	0·626	0·450
1·30	0·722	0·568
1·35	0·811	0·690
1·40	0·894	0·817
1·45	0·971	0·946

Strip, with grid

$\frac{r_c}{r_a} = 1{\cdot}71$ $(\beta^2)_a = 0{\cdot}444$

$\frac{r_c}{r}$	$\frac{r_c - r}{r_c - r_a}$	$\frac{V}{V_a}$
1·04	0·0927	0·0326
1·1	0·219	0·105
1·2	0·402	0·248
1·3	0·556	0·399
1·4	0·689	0·550
1·5	0·803	0·700
1·6	0·904	0·845
1·7	0·993	0·987

References

KLEEN, W. J., 1958, *Electronics of Microwave Tubes* (New York: Academic Press), p. 334.

LANGMUIR, I., and BLODGETT, K. B., 1923, *Phys. Rev.*, **22,** 347.

—— 1924, *Phys. Rev.*, **24,** 49.

MELTZER, B., 1959, *Brit. J. Appl. Phys.*, **10,** 391.

PIERCE, J. R., 1949, *Theory and Design of Electron Beams* (New York: D. Van Nostrand), p. 150.

SPANGENBERG, K. R., 1948, *Vacuum Tubes*, 1st edn (New York: McGraw-Hill), p. 462.

WATSON, E. E., 1927, *Phil. Mag.*, **3,** 849.

Note on the mechanical analogy of a viscoelastic fluid

The rheological equation for a viscoelastic fluid gives the response of the material to any imposed stress or strain. A certain amount of qualitative information may be derived from the study of mechanical methods or analogies which are designed to duplicate the observed time dependence of a real fluid; their behaviour is often more easily visualized than that of the corresponding fluid.

The models are generally made up of combinations of springs and dashpots, representing the elastic and viscous properties of the fluid respectively. The basic elements of the mechanical models, i.e. the Voigt element, a parallel combination of spring and dashpot, and the Maxwell element, a series combination, cannot by themselves demonstrate the non-linear relation between the applied shear stress and the shear rate, characteristic of real viscoelastic fluids (Fig. 1). It would be possible to model these by using a spring with a non-linear characteristic in the element, but the analogy is then too complex for demonstration purposes.

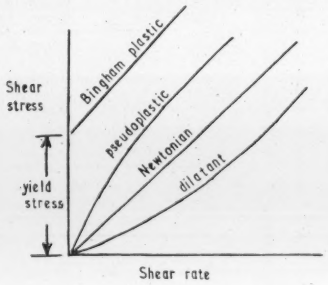


Fig. 1.

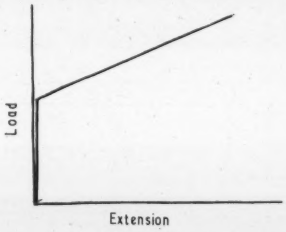


Fig. 2.

The Bingham plastic with its yield stress may be demonstrated by using a cylindrical spring wound with inter-turn pressure. With such a spring the adjacent coils exert a pressure between themselves when the spring is unloaded; the application of a tensile load reduces the inter-turn pressure, giving the load-extension characteristic of Fig. 2, which may be compared with the relationship for the Bingham plastic of Fig. 1.

Department of Engineering,
University of Manchester,
Manchester 13.

P. G. MORGAN
21*st April* 1961

Size distribution determinations of non-volatile droplets by light and electron microscopy

A method of replicating spread droplets originally described by the author (1959) necessitated the droplets being sampled on to a thin wet film of gelatin and therefore limited the method of sampling to either electrostatic precipitation or sedimentation. A non-volatile droplet of 10 μ diameter when spread out as a biconvex lens on the gelatin surface required a depth of gelatin of only 1 μ in order to be faithfully replicated. It has now been found that if the droplets are sampled on to a dry gelatin film it may be easily reconstituted to a sufficient depth to allow replication of the droplets to take place. Use of a dry film enables the application of the replicating technique to be extended to sampling by impaction and thermal precipitation.

The gelatin film is prepared by spreading a thin film of 2% gelatin–water solution on the cleaned sampling surface and allowing it to dry. As soon as the droplets have been sampled the deposit should be heavily breathed on. The surface of the gelatin is thereby reconstituted into a wet film of depth sufficient to allow the droplets, which spread to form plano convex lenses on the dry gelatin, to become biconvex lenses which will be replicated when the film dries again. This replicating stage is complete in a few seconds (compared with approximately 30 minutes for 'wet' sampling (Harris 1959)) and reduces to a minimum the possibility of any change in droplet diameter due to ageing. For electron microscopy the droplets are washed away and collodion and carbon replication carried out as previously described (Harris 1959). If light microscopy is to be undertaken it is advisable, but not essential, to wash away the droplets before direct examination by either lightfield or darkfield illumination (Fig. 1) or phase contrast. The last is desirable for high magnification examination. The replica diameters are converted to droplet diameters by applying the spread factor which is determined in the same way as that for a wet gelatin film (Harris 1959).

The reliability of the dry film method has been carefully checked with the wet film method by doing a series of simultaneous spread factor determinations for different droplet clouds and sizes and, in each instance, the values obtained have been in close agreement. Further confirmation of the accuracy of the method was obtained by sampling a droplet cloud on to dry gelatin film with the conifuge (Sawyer and Walton 1950). This instrument samples the droplets in such a way that they are deposited in a complete spectrum according to their settling velocities. Since in the present case all the droplets were of the same density, at any point of the deposit they were all the same size and, although they had spread, their true diameter was known. Hence in any

region of the deposit a direct check was provided for the values obtained by the dry gelatin method and agreement was again found to be close.

A valuable feature of the dry gelatin film technique is that it permits complete assessment of an aerosol consisting of

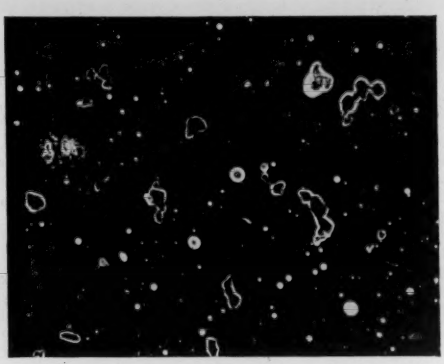


Fig. 1. Anthracite particles and oil droplets (darkfield illumination) ×250.

non-volatile droplets and particulate matter. The latter should not be readily water soluble otherwise errors may arise in the reconstitution stage. When electron microscopy of the sample is to be carried out, the particulate matter should in addition be insoluble both in benzene, which is used to wash away the droplets prior to collodion replication, and in the solvent of the collodion solution. The reconstitution stage will cause the particulate material to become partially embedded in the gelatin and, for electron microscopy, two ways of final processing are possible. If structural information is being sought, a collodion replica may be taken of both the droplet impressions and the exposed portion of the particles and dry stripped from the gelatin (Fig. 2). If size distribution determinations are to be made, the collodion replica is freed by dissolving the gelatin. Only the droplet

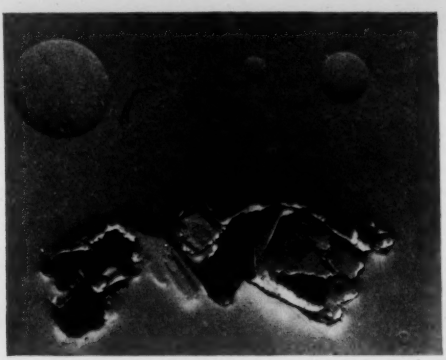


Fig. 2. Carbon replica of oil droplet impressions and anthracite particles. ×10 000.

impressions are replicated whilst the particles themselves adhere to the collodion.

The improvement in technique described was developed while the author was working at the Chemical Defence Experimental Establishment, Porton Down, Salisbury.

Experimental Station, Microbiological Research Establishment, Porton, Salisbury, Wilts.

W. J. HARRIS

25*th April* 1961

HARRIS, W. J., 1959, *Brit. J. Appl. Phys.*, **10**, 139.
SAWYER, K. F., and WALTON, W. H., 1950, *J. Sci. Instrum.*, **27**, 272.

Creep of cylinders in torsion

The creep of bodies in which the stress is not single valued is frequently of practical interest. Popov (1949) has presented an iterative method for calculating the creep of a beam under bending stresses, which takes transient creep into account, and is based on the known creep equation for single valued stresses.

As a method of testing isotropic materials when the temperature of the environment is difficult to control, angular deformation has the advantage that it is free from thermal expansion errors. This has led to a wide use of torsional creep as a laboratory method and the technique has been used with high sensitivity (Johnson 1950).

When the specimen cannot be in the form of a thin walled tube, as in Johnson's work, the problem arises which is the converse of that considered by Popov, namely, to what extent is the creep of a solid cylinder in torsion representative of creep at a single valued stress?

This note presents a simple analysis which shows that the observed deformation is closely similar to that under a single valued stress, even when the creep rate is highly stress dependent. The appropriate single valued stress appears from the analysis. A numerical example is given. It is apparent that tubes having comparatively thick walls give results of high accuracy, and that in many cases the accuracy obtained using solid specimens is adequate.

Symbols

$g(t)$, total angular deformation per unit length; α_e, elastic component of angular deformation; α_c, creep component of angular deformation; σ, shear stress; S, shear modulus; G, couple on cylinder; r, radius of an annular element of material; R, radius of cylinder; R_i, intermediate radius having stress at which creep under single valued stress is considered; $f = r/R_i$; $F = R/R_i$; A, B, a, b and n are constants of the material; $k = A(\sigma R_i)^b e^{-at}$.

Theory

It is required to evaluate $\dot{g}(t)$ in terms of $(\dot{\alpha}_c)_{R_i}$. The couple on the cylinder is

$$G = \int_0^R 2\pi r^2 \sigma dr$$

$$= \int_0^R 2\pi r^3 S \alpha_c dr$$

$$= \int_0^R 2\pi r^3 S(g(t) - \alpha_c) dr.$$

Consider creep under a constant couple, and make the approximation that S is constant. By inversion symmetry and the boundary conditions, $g(t)$ is independent of r. Thus, differentiating with respect to t,

$$\dot{g} = \frac{4}{R^4} \int_0^R r^3 \dot{\alpha}_c dr. \tag{1}$$

Now take a popular expression for the creep behaviour of a material, both transient and steady state (Finnie and Heller 1959),

$$\dot{\alpha}_c = \frac{B\sigma^n}{r}(1 + A\sigma^b e^{-at}).$$

The creep rate at the general radius r, in terms of the rate at the particular radius R_i, is then

$$(\dot{\alpha}_c)_r = f^{n-1}\left\{(\dot{\alpha}_c)_{R_i} + \frac{AB}{R_i}(\sigma R_i)^{n+b}(f^b - 1)e^{-at}\right\}$$

whence, by integration, Eqn (1) becomes

$$\dot{g} = (\dot{\alpha}_c)_{R_i} \cdot \frac{4}{F^4} \cdot \frac{1}{1+k}\left(\frac{F^{n+3}}{n+3} + \frac{kF^{n+b+3}}{n+b+3}\right). \tag{2}$$

As transient creep dies away and k approaches zero,

$$\dot{g} \to (\dot{\alpha}_c)_{R_i} \cdot \frac{4}{n+3} \cdot F^{n-1}.$$

This equation defines the appropriate value of R_i, since we wish to obtain $\dot{g} = (\dot{\alpha}_c)_{R_i}$. Thus

$$F^{n-1} = \frac{n+3}{4} \quad \text{or} \quad R_i = R\left(\frac{n+3}{4}\right)^{\frac{1}{n-1}}.$$

With this value of R_i, Eqn (2) becomes

$$\dot{g} = (\dot{\alpha}_c)_{R_i}\left[1 - \frac{k}{1+k}\left\{1 - \frac{4}{n+b+3}\left(\frac{n+3}{4}\right)^{\frac{n+b-1}{n-1}}\right\}\right]$$

i.e. the deformation is that observed under a single valued stress in so far as

$$1 \geqslant \frac{k}{1+k}\left\{1 - \frac{4}{n+b+3}\left(\frac{n+3}{4}\right)^{\frac{n+b-1}{n-1}}\right\}. \tag{3}$$

This inequality indicates the maximum difference between $\dot{g}$ and $(\dot{\alpha}_c)_{R_i}$ when k is large, which is at the commencement of creep. As k decreases the difference decreases. It is also apparent that $\dot{g} = (\dot{\alpha}_c)_{R_i}$ if the stress dependence of the steady and transient components of the creep is the same, for then $b = 0$.

A numerical example of interest to the author is provided by the data of Manjoine and Mudge (1954) on zirconium at 400° C and 6000 lb/in^2, for which $n = 8$ and $b = -2$. The other constants are $A = 7 \times 10^{11} \text{g}^2\,\text{cm}^{-4}$, $a = 5 \times 10^{-6}\,\text{s}^{-1}$, and $B = 4 \times 10^{-53}\,\text{gm}^{-8}\,\text{cm}^{16}\,\text{s}^{-1}$. These values give $\dot{g} = 0{\cdot}92(\dot{\alpha}_c)_{R_i}$ at $t = 0$ and $\dot{g} = 0{\cdot}99(\dot{\alpha}_c)_{R_i}$ at $t = 210$ h. Thus, though the creep rate varies as the eighth power of the stress, the observed rate of deformation never differs by as much as 10% from the creep rate appropriate to a single valued stress.

For a tube of similar material, $\dot{g}$ differs by less than 1% from $(\dot{\alpha}_c)_{R_i}$ at the commencement of creep if the wall thickness of the tube is less than $0{\cdot}14\,R_i$. For this wall thickness, R_i is 101% of the mean radius.

While similar analysis can be made using other creep equations for specific materials, these examples show that the existence of multivalued stress is not a severe handicap in torsional creep experiments.

Berkeley Nuclear Laboratories, R. V. Hesketh
Central Electricity Generating Board. 26*th April* 1961

At present on attachment to
United Kingdom Atomic Energy Authority,
Dounreay Experimental Reactor Establishment,
Thurso, Caithness.

Finnie, I., and Heller, W. R., 1959, *Creep of Engineering Materials*, 119 (New York: McGraw-Hill).
Johnson, A. E., 1950, *J. Sci. Instrum.*, **27**, 74.
Manjoine, M. J., and Mudge, W. L., 1954, *Proc. Amer. Soc. Test. Mater.*, **54**, 1050.
Popov, E. P., 1949, *J. Appl. Phys.*, **20**, 251.

New books

Magnetic materials. By F. BRAILSFORD. (London: Methuen; New York: John Wiley, 1960.) Pp. vii + 188. Price 16*s*.

The first version of Professor Brailsford's book on magnetic materials is well known and many of us have used it frequently. It has now been almost completely rewritten and much improved. Quite a lot of the very latest information has been included to bring the contents as up to date as possible.

The rationalized M.K.S. system of units is used throughout the book, and no concession is made to those of us who are old fashioned enough to think in gauss and oersteds. We will not even find these quoted by the side of the new units, as is so often done in recent books. Whilst welcoming the principle underlying this policy, it does make life sometimes a little difficult, and a little less austerity would have been very welcome.

The first chapter is called 'Introductory' but does not always seem to be sufficiently general. For instance, it goes into details of the mathematical expression of eddy current losses in laminations. The second chapter, however, on Ferromagnetic domains and properties of single crystals, gives what is to me an introductory and superficial description of the fundamental theories of magnetism. The above mild criticisms of positioning of subject-matter are the only ones I have to make.

The remainder of the book presents the information about magnetic materials which most users would require in a very clear and relevant form, and the physics, chemistry and often metallurgy underlying the various observed characteristics is well presented. With the inclusion of a chapter on ferrites, the book is certainly complete and in its present form covers the entire subject.

The diagrams are very informative and sufficiently accurate to be of quantitative use. On the whole, the book is sure to be most valuable for many students and engineers who want all the relevant information about magnetic materials in a clear and concise form at their fingertips.

K. HOSELITZ

Notes and comments

U.S. National Bureau of Standards

The U.S. National Bureau of Standards announces the publication of two new Handbooks—(i) Report of the International Commission on Radiological Units and Measurements (ICRU) 1959, Handbook 72, superseding the 1956 report which was published as Handbook 62; (ii) X-ray Protection up to Three Million Volts, Handbook 76, prepared by Sub-Committee 3 of the National Committee on Radiation Protection and Measurement. Both are available from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C., price 65 cents and 25 cents respectively.

Electro-Heat Congress

The British National Committee on Electro-Heat reports that the complete Proceedings of the 4th International Congress on Electro-Heat, held in Italy in May 1959, have now been published in two volumes. Volume 1 contains the programme and a general report of the Congress, all the contributions to the discussions and a complete index. Volume 2 includes in full all the 185 papers presented at the Congress. The papers range over the whole field of electric process heating, and are divided into ten sections: development of electro-heat in various countries, influence on productivity, statistics; electro-heat applications in the foundry; in ferrous and non-ferrous metal melting; in heat treatment, brazing, welding, etc.; in ceramics and glass; in wood, plastics, textiles, rubber and leather; in the foodstuffs, chemical and pharmaceutical industries; in electrochemistry and electrometallurgy; technical and economic problems involved in the development of electro-heat; problems of research and development techniques, temperature control, heat transfer, etc.

Publishers: Comitato Elettrotecnico Italiano, Via San Paolo 10, Milano, Italy. Price: 35,000 lire. Further information from the British National Committee, c/o British Electrical Development Association, 2 Savoy Hill, London, W.C.2.

Fourth Rubber Technology Conference, 1962

The Institution of the Rubber Industry is holding its Fourth International Rubber Technology Conference at the Church House, Westminster, London, from 22nd to 25th May 1962.

The main sections of the Conference will be: Latex science and technology; General rubber technology—equipment, processing, compounding and evaluation; Rubber science—chemistry, physics, testing, analysis; Materials and products—rubbers, textiles and ingredients.

The papers will be available to Conference members as pre-prints and each author will be given 10–15 minutes to summarize, in order to leave ample time for discussion. A large number of papers will be presented by U.K. delegates and present indications are that delegates from eleven other countries will also present papers.

Further details regarding papers, authors and officials will be issued in due course by the Institution of the Rubber Industry.

4th international symposium on gas chromatography

The 4th international symposium on gas chromatography, which is also the 41st meeting of the European Federation of Chemical Engineering, will be held at Hamburg from 13th to 16th June 1962. It is organized by the Analytical Chemical Division of the Gesellschaft Deutscher Chemiker and the Gas Chromatography Discussion Group of the Hydrocarbon Research Group of the Institute of Petroleum. The main language of the meeting will be English and the papers presented will be classified under three headings: Theory; Apparatus and Techniques; Applications.

Preliminary registration for attending the symposium should be sent as soon as possible to: Gesellschaft Deutscher Chemiker, c/o Dr. W. Fritsche, Frankfurt/M, Postfach 9075, Germany.

Conférence Internationale des Arts Chimiques

The Conférence Internationale des Arts Chimiques will take place from 25th April to 4th May 1962, in the halls of the Maison de la Chimie (Centre Marcelin-Berthelot), 28 bis, rue Saint-Dominique, Paris (7e). The VIe Salon International de la Chimie is taking place at the same time at the Centre National des Industries et des Techniques, Place de la Défense, Puteaux, Paris. This is an exhibition of the processes, techniques and plant employed in chemistry and its numerous industrial branches. Further details may be obtained from the Conférence Internationale des Arts Chimiques, 28, rue Saint-Dominique, Paris (7e).

School of Welding Technology

We have received details of courses to be held at the School of Welding Technology in October 1961 and also a provisional future programme for November to March.

The School of Welding Technology was established by the Institute of Welding in October 1957. Full time lecture courses, which may last up to one week or more, are held in London at the headquarters of the Institute and evening courses are held in London and in appropriate industrial centres. In general the courses are for engineers, designers, metallurgists and managerial and supervisory staff—in all cases for men who have received an engineering training.

Further information may be obtained from the Institute of Welding, 54 Princes Gate, Exhibition Road, London, S.W.7.

The Third International Congress on Catalysis

The Third International Congress on Catalysis is to be held in Amsterdam from 20th to 25th July 1964. The theme for the Congress will be The Mechanism of Heterogeneous Catalysis. The Secretary is Dr. D. M. Brouwer, c/o Badhuisweg 3, P.O. Box 3003, Amsterdam-N, Netherlands.

Notes on the preparation of papers

We will in future be publishing in each issue of the Journal notes on the preparation of papers to be submitted for publication. These instructions will give the main points to be considered. Full details are available in the Institute and Society's 'Notes for Authors', obtainable free of charge from the Editorial Office, 1 Lowther Gardens, Prince Consort Road, London, S.W.7.

Authors should submit two copies of their manuscript, typed in double spacing on one side of the paper, each with a set of small copies or prints of the diagrams attached. There should also be two extra copies of the title and abstract and a set of fair copies of the diagrams drawn in Indian ink with the lettering in soft or blue pencil. An important point to notice is the style of references which should be used. In the text references should be made by giving the name of the author and the year of publication, e.g. (Jones 1942), and details should be given in the list at the end where references are arranged in alphabetical order of authors' names.

British Journal of Applied Physics Sub-Committee

Members of the Sub-Committee for the British Journal of Applied Physics are as follows:

J. M. A. Leniham, M.Sc., Ph.D., A.M.I.E.E., F.Inst.P. (*Chairman*).

Prof. R. W. Ditchburn, M.A., B.Sc., Ph.D., F.Inst.P. (*Ex-officio member*).

W. Hirst, B.Sc., Ph.D., F.Inst.P.

C. A. Hogarth, B.Sc., Ph.D., F.Inst.P.

J. W. Menter, M.A., Ph.D., F.Inst.P.

A. C. G. Menzies, M.A., D.Sc., F.Inst.P.

Prof. D. A. Wright, D.Sc., F.R.A.S., F.Inst.P.

Journal of Scientific Instruments

Contents of the July issue

PAPERS

Improved fluorimeter for uranium analysis. By E. N. Haran.
Differential volumenometer for measuring minute volume changes in solids. By R. Lowrie and F. Slish.
Phase and modulation fluorometer. By J. B. Birks and D. J. Dyson.
Improved attachment for high temperature single-crystal x-ray work. By A. Barclay and J. D. Donaldson.
Fixed attenuators consisting of a network of identical resistors. By C. H. Vincent.
A method for gauge factor determination. By I. G. Scott.
Measurement of friction between rubber-like polymers and steel. By D. I. James.
Electrical micromanometer. By H. R. Hart.
Method of density determination on a micro-scale. By J. M. Jones.

RESEARCH NOTES

Device for the stabilization of helium cryostat temperatures. By C. J. Adkins.

LABORATORY AND WORKSHOP NOTES

A simple volumenometer. By R. G. Loasby.
Contactless resistivity meter for semiconductors. By J. C. Brice and P. Moore.
Loading device for x-ray cameras. By R. O. A. Hall.
Rubber as a sample for proton resonance fluxmeters. By S. J. Rogers.

NOTES AND NEWS

New Instruments, materials and tools New books Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.
EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.
REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.
ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611-2.)
SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12*s.* 6*d.* each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.
CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

PRINTED IN GREAT BRITAIN BY UNWIN BROTHERS LIMITED, WOKING AND LONDON

Summarized Proceedings of a Symposium on Electronic Devices at Helium Temperatures—London, November 1960

Abstract

A one-day symposium on electronic devices which work at helium temperatures was held jointly by the Low Temperature Group and the Electronics Group of The Institute of Physics and The Physical Society on 15th November, 1960. The emphasis was on devices which could be used in electronic computers. Recently there have been considerable advances in the research on Crowe cells and cryotrons. Both devices can be switched at high speed (~ 10–15 ns) and the factors governing their performance are fairly well understood. There have been equal advances in the development of a helium refrigerator of great reliability and efficiency. With cryotrons, the conditions for reproducibility are far less stringent than with the Crowe cell and, as four terminal devices with current gain, they offer great flexibility to computer designers, being capable of use both in the store and in logical components. The p–n junction, tunnel diode, and cryosar are all based on semiconductors and all work at helium temperatures. They could be very useful ancillary devices in computers based on superconductors or, on the other hand, computers could be based on these devices entirely and the use of liquid helium is then unnecessary. However, taking into account the low heat dissipation and close packing feasible with superconductive devices (10^7 bits in a 35 cm cube), where fast computing is concerned the cryotron offers great advantages. It is, moreover, a device of universal application throughout a computer which may demand a new approach to logical design and the layout of the machine.

Introduction

EIGHT papers were read at this symposium; they were 'An introductory survey' by Dr. D. H. Parkinson (Royal Radar Establishment), 'Superconductors' by Dr. J. M. Lock (Royal Radar Establishment), 'Crowe Cells' by Dr. E. H. Rhoderick (Services Electronics Research Laboratories), 'Cryotrons' by Dr. D. R. Young (I.B.M., Poughkeepsie, U.S.A.), 'Work on cryotrons at the N.P.L.' by Dr. J. S. Hill (National Physical Laboratory), 'p–n junctions at low temperatures' by Dr. A. K. Jonscher (G.E.C. Research Laboratories), 'Computers of the future' by Dr. J. S. Gill (Ferranti Ltd.), and 'The cost of using helium devices' by Dr. E. Mendoza (Manchester University). This report summarizes their contents and the subsequent discussions which tended to centre on the use of devices in computers.

The first part of this report deals with superconductors and devices based on them, the second part with low temperature devices based on semiconductors, the third with liquid helium refrigeration and consequent problems, finally some appreciation of the future outlook is given.

Devices based on superconductors

From the time when superconductivity was first discovered in 1911 thoughts have turned to possible applications of the phenomenon of zero resistivity. The first ideas concerned the generation of high magnetic fields in superconducting solenoids which has received renewed attention recently. Since then there have been many applications of superconductivity most of which have remained inside low temperature laboratories. The superconducting bolometer was developed during the war by Andrews as a detector of infrared radiation and has continued to receive attention. Other examples are superconductive amplifiers of very high sensitivity (e.g. de Vroomen and Van Baarle 1957), the superconducting galvanometer (Pippard and Pullan 1952) capable of detecting 10^{-5} ampere in a resistance of 10^{-7} ohm, besides various switching devices (e.g. Templeton 1960).

The cryotron was first described by Buck (1956) and was soon followed by the Crowe cell (Crowe 1957). These two devices and particularly the cryotron have received much attention. Two rather similar devices, the persistatron (Buckingham 1957), and the persistor (Crittenden 1957) have also been proposed though there has been little recent work on them. The Crowe cell and cryotron work in quite different ways; to appreciate their operation and the significant differences, it is first necessary to outline the relevant features of the superconducting transition itself.

A superconductor is a metal such that below a well-defined temperature, the transition temperature T_c, all traces of electrical resistivity are non-existent. This sharp transition is not only characteristic of pure metals, but also impure metals and alloys provided they are sufficiently ordered. Thus well annealed tin specimens containing 2% of indium can have transitions only about one thousandth of a degree wide, even though such an impurity concentration is sufficient to increase the normal resistivity just above the transition to 1000 times that of the purest tin. This feature is significant for superconducting devices where a reasonably large normal resistance is required. Roughly one-third of all the metallic elements are superconductors above about 0·1° K. None of the monovalent metals or alkaline earth metals are superconductors. There are many superconducting alloys some of which have components which are themselves not superconductors.

In the superconducting state the resistance of the material is zero. The most sensitive test of this point is provided by the Crowe cell itself in which the persistent current shows no observable decay over long periods of time. It can easily be shown that the resistance in the superconducting state must be less than 4×10^{-14} of its normal value at the most. Below T_c, the normal resistance can be restored by applying a high enough external magnetic field, the critical field H_c, which plotted as a function of temperature is parabolic (see Shoenberg 1952). The destruction of superconductivity by an applied field can be sharp, e.g. 0·1 gauss wide, and forms the basis of operation for the cryotron.

It follows from the property of zero resistance that in the superconductive state the value of B, the magnetic flux, in the material cannot change if the external field is changed. Eddy currents induced on the surface of the metal just compensate for any change of external field; such currents are

'super-currents' and flow without decay. However, the interesting feature of superconductors is that not only is dB/dt zero but so also is the value of B. Thus if a specimen is cooled in a magnetic field from above the transition, the flux is suddenly pushed out completely as it passes through the transition temperature. This is the Meissner effect shown first by Meissner and Ochsenfeld (1933). There is a similar effect if the external field is slowly reduced through the critical value H_c at constant temperature.

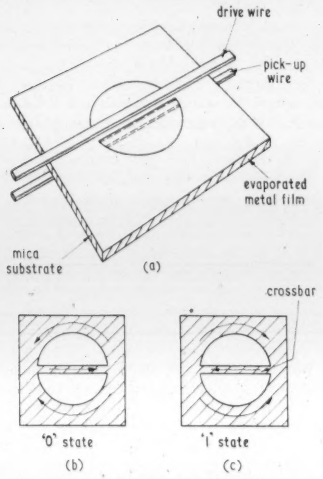


Fig. 1. Diagram of Crowe cell.

When the material is superconducting, the state with $B = 0$ can be regarded as that of thermodynamic equilibrium, and the critical field curve can be treated as the equilibrium curve between the normal and superconducting phases. One can then use simple thermodynamic reasoning to show that the superconducting phase always has lower entropy than the normal phase and is thus more ordered.

The surface currents, responsible for the expulsion of the magnetic flux from a superconductor flow in a surface layer, typically about 1000 Å thick, which extends to the depth to which magnetic fields penetrate, the penetration depth δ. For specimens of thickness similar to δ, such as the films used in Crowe cells and cryotrons, there are two immediate consequences. First the diamagnetic susceptibility falls below the bulk value $1/4\pi$, from which it can be shown that the critical field is considerably increased.

The critical current is also affected in such thin specimens. In bulk specimens the critical current is just sufficient to produce the critical field at the surface of the wire. For films thinner than the penetration depth the critical current is less than that required to produce the critical field at the surface. This is in contrast to the increased critical field and can be shown to be due to the form of the magnetic field and related current distribution through the thickness of the film.

This brief outline of the features of the superconductive transition is adequate for the consideration of the Crowe cell and cryotron.

The Crowe cell consists of thin tin specimens usually evaporated on to mica in which there are two D-shaped holes, back to back, Fig. 1(*a*). It is usually made by leaving a clear circle in a first evaporation and then depositing the 'cross-bar' in a second evaporation. The two states of a cell (0 or 1) are represented by the sense of the persistent supercurrents which can be made to flow round the D-shaped holes as in Figs 1(*b*) and 1(*c*). Drive and sense wires are arranged above and below the cell. The whole assembly may be made by evaporation on to a single substrate, with insulation between successive conductors being effected by evaporated layers of silicon monoxide or other suitable material. Alternatively the film with the D's may be put down first on one side of a mica sheet with the drive wire on the other. The sense wire may then be evaporated on to a second sheet which is then accurately aligned with the first.

The mode of operation can be demonstrated in terms of Fig. 2 where the current induced in the cross-bar is shown corresponding to drive current pulses. The induced currents flow in the cross-bar and film so as to oppose any penetration of flux into the film, i.e. through the D's. Assuming unity coupling, during the rising edge of a drive pulse, AB, the cross-bar current, PQS, rises until it reaches the critical value at Q (determined by the thickness of the film), when the flux penetrates. During the rest of the rising edge a pulse is generated in the sense wire. At Q the cross-bar becomes normal and heat is generated in it. When the drive pulse reaches a constant value, the cross-bar current falls according to the L/R time constant of the cross-bar and the

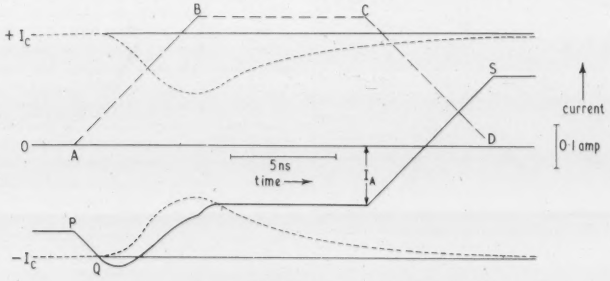


Fig. 2. Switching of Crowe cell.
——— crossbar current, – – – – – – critical current,
— — — drive current.

return path through the film, and eventually reaches the critical value again. The value of the critical current will itself have been reduced owing to the heat generated in the cross-bar. The peak BC of the drive current should be longer than the time constant L/R and the time constant for the dissipation of heat to the substrate. When the drive pulse is removed CD, the cross-bar current, reflects the change and a supercurrent is left circulating in the cross-bar. The operation with drive pulses of opposite sign, half pulses and repeated pulses of the same sign follows logically.

Typical Crowe cells are about 2 mm in diameter with cross-bars about 0·2 mm wide and 1000 to 2000 Å thick. They have been made successfully with both lead and tin evaporated on to both glass and mica. The fastest operation has been with those using tin evaporated on to mica, where pulses as short as a few nanoseconds have switched cells with cross-bars of the order of 900 Å thick. In general Crowe cells can be switched with pulses 10–15 ns long and with current amplitudes of about 100 mA. The corresponding sense wire signals are about 10 mv. It has been shown that information can be stored for long periods of time and remain intact against many half-level pulses. Thus the cell is suitable for use with a coincident current drive system.

The Crowe cell is a passive device and the output from one Crowe cell cannot drive another. In any computer with a store based on this cell, the logic would have to be done using other devices. To utilize the high switching speed, the read and write circuits must be arranged so that wires cross the store once only. Fortunately, owing to the small size of the cell, some 10^7 bits could be contained in a 35 cm cube, and the transit time for a pulse to reach the store, cross it once and return to the generator is easily made compatible with the switching speed. A coincident current selection system for writing and reading has been devised, based on the cryosar. The sense signals have also been shown to be capable of switching Esaki diodes (see below) from which a read out register would be constructed. These systems have been described by Lock, Parkinson and Roberts (1961).

It has also been shown that heat dissipation in the store is not a critical problem; a repeated switching rate of 50 Mc/s seems possible. The mechanism of the return of normal resistance in the cross-bar has been studied to some extent. Thermal propagation of the normal phase along the cross-bar plays some part and has at least led to the development of another bistable device, the calotron, which had been described by Broom and Rhoderick (1960); whether this could be used in computers is doubtful.

The Crowe cell, however, suffers from one serious drawback, the tolerances which must be set on the critical current of the cross-bar, on the magnitude of the stored current and on the drive currents when used in coincident current manner, are very tight, and at the best are about 7% to 8% for each (Peacock 1961). This feature may well prevent the Crowe cell having any useful future.

In its original form the cryotron consisted of a straight superconducting wire around which a second superconductor with a higher transition temperature was wound as a small solenoid. The field generated by the solenoid was sufficient to destroy superconductivity in the straight wire, but not in the solenoid itself. Such a device could be used rather as a simple Post Office relay, but is obviously far too slow for use in a modern computer. The cryotron has been developed to consist simply of two thin strips of different superconductors of different width crossing each other at right angles, as shown in Fig. 1 of the accompanying article by Dr. D. R. Young (this will be referred to again as (D.R.Y.)) One strip of width W_c, is the control wire and is usually of lead, while the other of width W_g, is the gate and is usually of tin; W_c is considerably less than W_g. Such an arrangement would usually be made by evaporating the strips on to a glass substrate, but interleaved with an evaporated insulating layer of silicon monoxide. In use, it is operated just below the transition temperature of tin (e.g. 2·9° K) which is far below the transition temperature of lead. The tin gate remains superconducting for currents below its critical current I_c. The strip immediately below the control wire can be made normal if a current through the control wire is sufficient to create the critical field H_c of the gate. Suppose the smallest current to do this is i. Then if i is less than I_c the cryotron has current gain $g = I_c/i$ or a supercurrent I_c in a circuit can be switched by a control current i. Now ignoring in the first place any effect of the ground plane, $H_c = \alpha i/W_c$ where α is a constant and the gate current $I_c = i_c W_g$, where i_c is the critical current per unit width.

Hence we can write

$$g = \frac{\alpha i_c W_g}{H_c W_c} = \frac{H_i}{H_c}\frac{W_g}{W_c}$$

where H_i is the surface field produced by the current i_c per unit width in the gate. We can also write $g = f(t, T)W_g/W_c$, where t is the thickness of the film and T is the temperature of operation. If the film thickness is reduced, f decreases and so g is impaired. On the other hand, when the gate is driven normal the largest possible return of resistance is required which implies the use of the thinnest possible films. These requirements are conflicting and there is an optimum thickness of about 3500 Å. As originally made the current gain of the crossed strip cryotrons was very far short of the theoretically possible values. There have been two factors which have completely changed this situation. First the use of a ground plane (Young 1959), and second the elimination of edge effects in the gate (Delano 1960).

When a strip conductor runs parallel to a superconducting ground plane at a distance small compared with its width the image, in the form of supercurrents in the ground plane, modifies the current distribution and the field between the strip and plane. Between the strip and plane the field is virtually doubled and made more uniform, elsewhere the field is reduced to zero. The gate current is made more uniform across its width and consequently the gate critical current is increased. Likewise the resistance restored in the gate is virtually confined to a strip of length W_c.

Edge effects arise with evaporated films because there is always some sort of penumbra at the edge where the film gets thinner and where there can be many defects. Imperfections change the transition temperature and a change in thickness will change the critical field, both leading to ill-defined transitions. In fact, removing the edges of a film by mechanically trimming it produces a very well-defined and sharp magnetic transition at a lower field. Similarly the temperature transition is also sharpened. Trimming the edges does not affect the critical current appreciably. Annealing, in which the grain size in the films is made more uniform, also improves the transitions. Reduction of the critical field obviously increases the current gain.

Cryotrons which perform in a manner near to the theoretical can now be made. Removal of the edge effects also greatly improves the reproducibility. Gains as high as 10 have so far been achieved though 3 or 4 is quite enough for most purposes. The time constant for switching L/R is governed by the inductance L of the circuit loop involved and the value of the restored resistance R. Time constants

as small as approximately 10 ns have been observed. Where the highest speed is required the restored resistance must be high and hence ω must be large; in consequence current gain has to be sacrificed. Control currents of the order of 500 mA are adequate, depending on how close the operating temperature is to the superconducting transition in tin.

The time constant can be shortened by increasing the restored resistance which also reduced the mismatch to a typical superconducting strip transmission line. To this end the 'in-line' cryotron has been studied recently. Here the control and gate wires are parallel and the section of gate which is made normal is very much larger. This arrangement is discussed in the accompanying article (D.R.Y.).

Cryotrons can be used in digital computers in a number of ways; the following are three examples. First, they may be connected as cross-coupled flip–flops as was originally proposed. Second, they can be used as simple switches where a supercurrent is switched into one of two paths, each path being sensed by another cryotron. Third, Haynes (1960) has described an ingenious way in which the information is stored as a circulating current in a circuit loop, Fig. 3. A, B

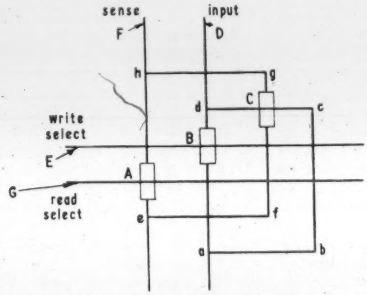


Fig. 3. Cryotron persistent current store.

and C are cryotrons; a persistent current in the loop abcd represents 1 and no current 0. To set up a 1, current is applied to wire D and a pulse to E, the gate of B becomes resistive and the current in D is diverted through abcd. Removal of the current from E leaves the current in D still flowing through abcd. The external current supply to D is then cut off leaving a circulating current in the loop abcd. Sensing a 0 or 1 can be carried out by applying a current to wire F, the sense line, and also a pulse to the 'read select line' G. Current in G makes the gate A resistive, and if there is no circulating current in abcd the current in F is diverted through the path efgh which contains the gate of cryotron C. If there is a circulating current in abcd, the current in F cannot flow and is diverted to a parallel superconducting line not shown, so controlling a digit register cryotron which will record the stored 1. Units such as in Fig. 8 are repeated in square array throughout the store. Reading and writing simultaneously but on different parts of the store is possible; reading is also non-destructive.

Tolerances on cryotrons are not at all severe, provided the current gain is greater than 1.

Devices based on semiconductors

The devices of immediate interest are: the p–n junction of which the Esaki or tunnel diode is a particular case, and the cryosar. It ought to be possible to make transistors work satisfactorily in liquid helium. In the accompanying paper by Dr. A. K. Jonscher, there is given the general pattern of behaviour of germanium and silicon p–n junctions as the temperature is lowered. The Esaki or tunnel diode (Esaki 1958) is a particular case of the p–n junction where the level of doping in the two halves is very high. In the p-type material enough acceptors are present to depress the Fermi level into the valence band while in the n-type there are enough donors to raise the Fermi level into the conduction band, Fig. 4(*a*). Suppose the overlap of the band edge is Δ.

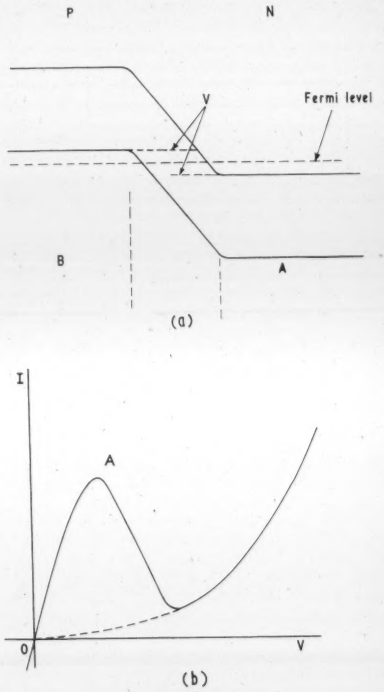


Fig. 4. Esaki diode characteristics.

With no voltage applied across the junction electrons can pass from the Fermi level in A to that in B and vice versa by tunnelling across the junction. This is non-classical process (see, for example, Purcel 1960) and for it to be appreciable the junction must be very narrow. With a small voltage applied across the junction in either direction a current will flow. However, when voltage $V = \Delta/e$ is applied in the forward direction the band overlap will be destroyed and the currents will tend to zero. For greater voltages the overlap is completely destroyed, the situation is comparable with a normal p–n junction with the corresponding current–voltage characteristic. The important part of the characteristic is the region of negative resistance, separating the two parts of the characteristic representing stability and in principle these two states could be used to represent a 0 or 1 in a computer circuit. They represent two different impedances, conse-

quently the circuits would have to be arranged to sense the impedance of the device. By biassing an Esaki diode to a point such as A in Fig. 4(*b*) a very small voltage pulse, such as that from a Crowe cell, is required to drive it to the second part of its characteristic. This has been demonstrated and hence Esaki diodes could be the basis of an output register for a store composed of Crowe cells.

The cryosar (McWhorter and Rediker 1959) is a simpler device relying on the avalanche effect in a semiconductor. Suppose a sample of germanium which contains n-type impurities is cooled and a potential is applied across it. At low enough temperatures phonon scattering of the current carriers is negligible and impurities remain as the chief source of scattering. Provided the concentration of impurity is not too great then between collisions current carriers can absorb enough energy from the potential field to ionize impurity atoms on collision. Thus when the field is high enough the number of carriers increases very rapidly with time and avalanching of carriers occurs. With germanium at helium temperatures a symmetrical current–voltage characteristic is produced as in Fig. 5. The 'corners' such as at X

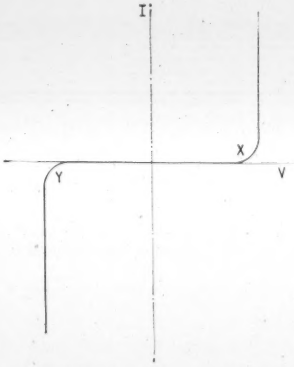


Fig. 5. Cyrosar characteristics.

and Y can be fairly sharp; compare with Fig. 3 of the paper by Dr. A. K. Jonscher.

Helium refrigeration

There is a natural reluctance with computer engineers to consider the computer based on cryogenic devices using liquid helium. The first good reason for this is that the refrigerating device must be utterly reliable and must run reliably for many thousands of hours continuously (1 year = 8760 hours). Laboratory helium liquefiers have not been sufficiently reliable by two orders of magnitude. A second reason is that building computer circuits at room temperature is bad enough, for many circuits have to be adjusted finally by trial and error, to have to cool to helium temperatures during building would be an almost insufferable difficulty. Once built, the computer itself must be reliable and withstand repeated coolings. 'Trouble shooting' at a late stage would also be very difficult. There are replies to all these objections.

Recently the reliability of helium refrigerators has increased markedly. A first prototype machine designed by A. D. Little for continuous closed circuit operation has 1000 hours operation to its credit without any adjustments whatever. (A later report suggests that wear in the driving compressor stopped operations at approximately 2500 hours.) This is within an order of magnitude of what is required. The machine has been described by MacMahon and Gifford (1960).

The principle of operation can be seen by reference to Fig. 6; there R is a regenerator connected as shown to a

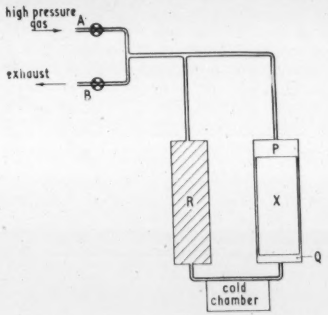


Fig. 6. Principle of helium refrigerator.

cylinder containing a displacer X. A and B are inlet and outlet valves respectively, both at room temperature. To begin the cycle imagine the lower end of R cold, the system at 1 atmosphere, and B closed. When A is opened to the high pressure source gas at approximately 300 lb/in^2, gas enters the space P above X and fills the refrigerator. Gas reaching the lower end of R is cooled. When pressure has been built up, X moves upwards, gas is displaced from P through R to Q and is cooled in R. Extra high pressure gas is drawn in through A during this process. A is now closed and B opened. The high pressure gas exhausts through R, cooling it in so doing and carrying heat from the low temperature chamber (Joule effect, and compare Simon expansion liquefier). The displacer finally moves downward and the last cold gas passes out from Q through R. B can now be closed and the cycle repeated. This system is not as efficient thermodynamically as the normal expansion engine; this is offset, however, by the use of a regenerator instead of heat exchangers. Moreover, the only moving part which is cold is the displacer which does not do any work and the only valves are at room temperature. It is easy to see that the regenerator could be placed inside the displacer itself, so combining the cylinder and regenerator unit. For the production of helium temperatures two models of a three-stage machine have been developed, the larger has a refrigerated space of about 1 foot cube and the smaller a space of about four inches cube. Obviously this machine is in its early stages of development and the present promising results show that the first objection to the use of helium can be removed.

The second objection arises from thinking in terms of conventional computers in which there are many different types of circuit element. Computing circuits based on cryotrons are very suitable for design by computers and contain fewer individual components. Complete circuits would have to be designed and produced in one set of evaporations. In future computers of large size and speed it will be necessary to pack circuit elements in large numbers into very small volumes. According to the devices used, techniques of 'microminia-

turization' and 'solid circuitry' or sheets deposited by evaporation will have to be used. Whether cryotrons are used or not we are soon to be faced with the necessity of producing complete circuits which work properly at once and need no adjustment. 'Trouble shooting' and replacement of sections of the computer even when cold are not insufferable difficulties provided the computer is designed with these things in view. Remote handling techniques for dealing with circuits at helium temperatures are quite feasible.

As far as cost is concerned, from what is known of the present complete Collins type of liquefiers it is likely that $50 000 should cover a complete installation to deal with a 10^7 bit store. This is not very different from the cost of an air conditioning plant to keep a large modern computer of similar storage capacity at room temperature.

The future outlook

The use of electronic computers is so widespread that it would be difficult to list all the many types of problem to which they are applied. There is no doubt whatever that eventually computers will be used to control all the major industrial processes and most of the minor ones too, most transport problems, traffic by air, sea and land, communications, language translations; besides the more obvious design problems. The field of application is so vast that the electronic computer will, and is indeed having, an impact on our life which is far greater than that of nuclear energy. Unfortunately, this is happening with less publicity and far fewer people are aware of it.

It is likely that large central computers having been set-up will themselves have to be interlinked to exchange information when necessary. This will magnify time of flight problems and consequently computers will have to be capable of dealing with several problems at once, 'switching attention' from one to the other while waiting for data to come in or for the answers to parts of the problem.

Of the devices which have been discussed the cryotron stands out as a device which commends itself for use in fast computing. If it were not for the need to use helium its future would be completely assured; meanwhile since this symposium it is now reported that I.B.M. hope to have a computer based on this device working in 1963.

Ministry of Aviation,
Royal Radar Establishment,
St. Andrews Road,
Great Malvern,
Worcs.

D. H. PARKINSON

References

BROOM, R. F., and RHODERICK, E. H., 1960, *Solid State Electronics*, **1**, 314.

BUCK, D. A., 1956, *Proc. Inst. Rad. Engrs, N.Y.*, **44**, 482.

BUCKINGHAM, M. J., 1957, *Low Temperature Physics and Chemistry*, *Proc. 5th Int. Conf.*, p. 229 (Univ. Wisconsin Press).

CRITTENDEN, E. C., 1957, *Low Temperature Physics and Chemistry*, *Proc. 5th Int. Conf.*, p. 232 (Univ. Wisconsin Press).

CROWE, J. W., 1957, *I.B.M. Journal*, **1**, 195.

DELANO, R. B., 1960, *Solid State Electronics*, **1**, 381.

ESAKI, L., 1958, *Phys. Rev.*, **109**, 603.

HAYNES, M. K., 1960, *Solid State Electronics*, **1**, 399.

LOCK, J. M., PARKINSON, D. H., and ROBERTS, L. M., 1961, *Proc. Inst. Elect. Engrs*, in the press.

MACMAHON, H. O., and GIFFORD, W. E., 1960, *Solid State Electronics*, **1**, 273.

MCWHORTER, A. L., and REDIKER, R. H., 1959, *Proc. Inst. Rad. Engrs, N.Y.*, **47**, 1207.

MEISSNER, W., and OCHSENFELD, R., 1933, *Naturwissenschaften*, **21**, 787.

PEACOCK, 1961, *Cryogenics*, in the press.

PIPPARD, A. B., and PULLAN, G. T., 1952, *Proc. Camb. Phil. Soc.*, **48**, 188.

PURCEL, R. A., 1960, *Solid State Electronics*, **1**, 22.

SHOENBERG, D., 1952, *Superconductivity*, p. 9 (Cambridge: University Press).

TEMPLETON, I. M., 1960, *Solid State Electronics*, **1**, 258.

DE VROOMEN, A. R., and VAN BAARLE, C., 1957, *Physica 's Grad.*, **23**, 785.

YOUNG, D. R., 1959, *Progress in Cryogenics*, **1**, 1.

Recent developments in high speed superconducting devices

by D. R. YOUNG, Ph.D., International Business Machines Corporation, Yorktown, New York

MS. received 4th April 1961

Abstract

The factors controlling the speed of useful cryotron circuits are: circuit geometry, insulation thickness, resistivity of superconducting film in normal state, superconducting and transition properties of films used. These factors are discussed in detail and in addition new circuit techniques are described that increase the speed of operation to the millimicrosecond range. A new type of cryotron is mentioned that introduces a significantly larger resistance in the normal state. This device is particularly useful for driving long lines such as encountered in cryogenic memory planes.

Introduction

THE operating speed of wire-wound cryotrons investigated by Buck (1956) was limited by the L/R time constant where L is the total loop inductance and R the resistance of one element in the normal state. The use of thin film techniques has reduced this time constant from the order of 10^{-3} second to the order of 10^{-7} second. However, the L/R time constant remains as the chief limitation in speed of operation. New techniques for reducing this to the millimicrosecond range will be described.

Operating speeds of simple circuits

The construction of a simple cross film type cryotron is shown in Fig. 1.

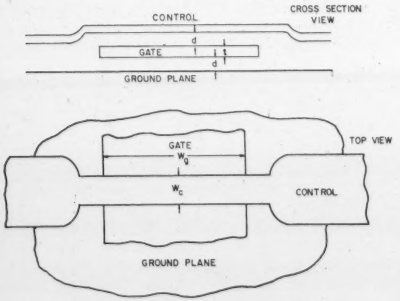


Fig. 1. Crossed film cryotron.

For the simple cryotron shown the inductance is given by*

$$L_{\text{control}} = \mu_0 \frac{W_g}{W_c} d\left(2 + \frac{\lambda_c + \lambda_{GP} + 2\lambda_g}{d}\right) \quad (1)$$

* For a discussion of techniques for computing inductance taking penetration effects into account, see Young *et al.* (1960).

where $\mu_0 = 4\pi \times 10^{-7}$, λ_c = penetration depth in control, λ_{GP} = penetration depth in ground plane and λ_g = penetration depth for gate. The resistance is given by

$$R = R_0 W_c / W_g \quad (2)$$

where R_0 is the film resistivity in ohms/square. Therefore,

$$\frac{L_c}{R_g} = \frac{\mu_0}{R_0} d\left(2 + \frac{\lambda_c + \lambda_{GP} + 2\lambda_g}{d}\right)\left(\frac{W_g}{W_c}\right)^2. \quad (3)$$

If $\lambda_{GP} = \lambda_c = 500$ Å, $\lambda_g = 1500$ Å,† $d = 5000$ Å and $R_0 = 4 \times 10^{-3}$ ohm/square, then

$$L_c/R_g = 4{\cdot}4 \times 10^{-10}(W_g/W_c)^2. \quad (4)$$

The film resistivity R_0 is determined by the gate film material and by the thickness. For tin the thickness dependence is shown in Fig. 2. Lead is used both for the control

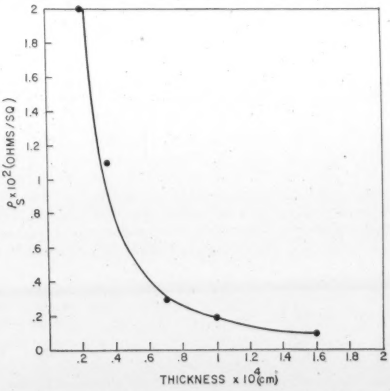


Fig. 2. Resistivity as a function of thickness for thin film.

film and the ground plane. The penetration depth in lead at the operating temperature is 500 Å.

Satisfactory circuit operation requires that the cryotron be able to control more current than is needed to do the controlling. This ratio is defined as the gain i_{g0}/i_{c0} which is given by $i_{g0}/i_{c0} = f(t, T)W_g/W_c$ where t is the film thickness and T is the temperature. Data showing the function $f(t, T)$ are given in Fig. 3. The resistance increases as the thickness is decreased whereas the gain decreases so there is an optimum thickness for maximum speed. For tin this is about 3500 Å, however, the energy dissipated by the circuits during transfer

† The reason for this will be discussed later.

processes is $\frac{1}{2} LI^2$, the energy stored in the magnetic field, which decreases as the operating temperature approaches the critical temperature. For this reason a film 6000 Å thick

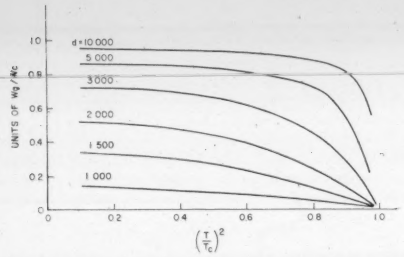


Fig. 3. Gain of crossed film cryotron as a function of reduced temperature.

was chosen for the calculations presented in this paper. If a gain of 2 is desired the width ratio required is 5. For a more complete discussion of the effect of gain on circuit operation see Young (1959).

In a practical circuit the inductance of the interconnecting lines must be taken into account and in addition it is generally necessary to have one input cryotron drive several output cryotrons with the net result that the time constant is considerably longer than that given by Eqn 3. For the two-

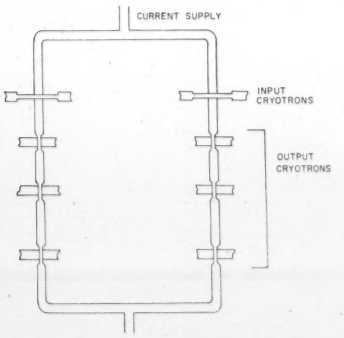


Fig. 4. Two-input–six-output loop for use in computing.

input–six-output loop as shown in Fig. 4 the time constant is

$$\frac{L}{R} = \frac{\mu_0}{R_0}\frac{dl}{W_c}\left(1 + \frac{2\lambda_{GP}}{d}\right) + \frac{12\mu_0}{R_0}\left(\frac{W_g}{W_c}\right)^2 d\left(1 + \frac{\lambda_{GP} + \lambda_g}{d}\right) \quad (5)$$

where l = length of the loop including the connecting lines, control and ground plane are all made out of the same material. For large W_g/W_c the time constant is determined by the devices themselves (i.e. the second term), taking account of the fact that there are 6 control lines rather than one, where as for small W_g/W_c the inductance of the connecting lines is important (i.e. the first term). These results are shown in Fig. 5.

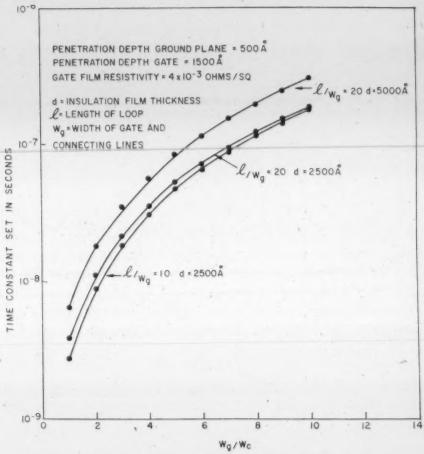

PENETRATION DEPTH GROUND PLANE = 500 Å
PENETRATION DEPTH GATE = 1500 Å
GATE FILM RESISTIVITY = 4 x 10⁻³ OHMS/SQ
d = INSULATION FILM THICKNESS
ℓ = LENGTH OF LOOP
Wg = WIDTH OF GATE AND CONNECTING LINES
ℓ/Wg = 20 d = 5000Å
ℓ/Wg = 20 d = 2500Å
ℓ/Wg = 10 d = 2500Å
TIME CONSTANT SET IN SECONDS
Wg/Wc

Fig. 5. L/R time constant as a function of width ratio.

Use of biased structures

The strong dependence of L/R on W_g/W_c leads one to do everything possible to reduce W_g/W_c. One way to do this is to use biasing, taking advantage of the fact that the i_g–i_c characteristics are curved and that a small signal gain can be greater than one even though the overall gain is less than one. This is illustrated schematically by Fig. 6. The circuit

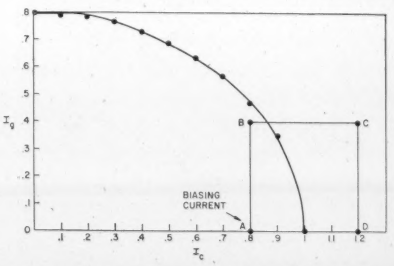


Fig. 6. Gain characteristic for cryotrons illustrating the use of biasing.

operates within the square ABCD. The initial starting point for one side of the loop is B whereas for the other side it is D. Data for actual cryotrons show that W_g/W_c can be reduced to at least 2 and still obtain satisfactory operation with biasing. This brings the L/R time constant of simple loops into the millimicrosecond region. The biasing current can be supplied to an additional film placed on top of the control film.

The use of biasing also permits the possibility of an in-line cryotron (see Fig. 7). The overall gain of this device is always less than one whereas small signal gains of ten have been observed. The operation of this device depends on the sign of the control current with respect to the gate current.

The characteristic to be expected using a critical field hypothesis is shown in Fig. 7. This predicts an overall gain of 1 and an infinite small signal gain, for the antiparallel case. It

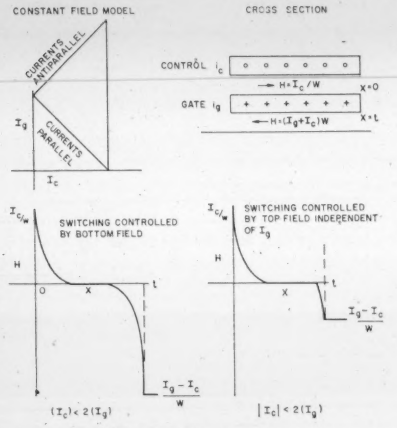


Fig. 7. In-line cryotron.

is to be noted that $i_{g\,max}$ occurs for $i_g = 2i_c$ also that for $i_g = i_c$ the field distribution is similar but reversed to that for $i_c = 0$. These features are independent of the model used. The infinite gain occurs due to the fact that the magnetic field on top of the control film is determined by i_c alone and governs the transition when it is larger than the field on the bottom of the film. For $i_g = 2i_c$ the fields on top and bottom are equal and for $i_g = i_c$ the field on the bottom is zero. Actual measured characteristics become more like those described above as the film thickness t becomes larger with respect to the superconducting transition depth λ.

A better model results by assuming that the transition occurs when the current density exceeds a critical value. The characteristics predicted by this model are shown in Fig. 8.

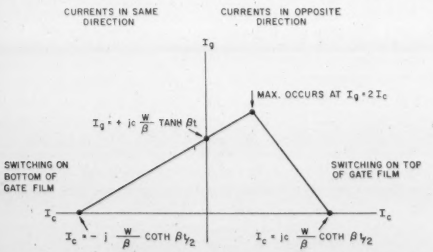

CURRENTS IN SAME DIRECTION
CURRENTS IN OPPOSITE DIRECTION
MAX. OCCURS AT $I_g=2I_c$
SWITCHING ON BOTTOM OF GATE FILM
SWITCHING ON TOP OF GATE FILM

Fig. 8. Gain characteristic of in-line cryotron using critical current 'density hypothesis'.

The overall gain is less than 1 and approaches 1 with increasing $\beta t(\beta = 1/\lambda)$. The maximum still occurs for $i_g = 2i_c$ and i_g for $i_g = i_c$ is equal to i_g for $i_c = 0$. The maximum small signal gain is not infinite. This arises from the fact that although the field on top of the gate film is independent of the gate current the slope and hence the current density is affected. This is shown by the field plot of Fig. 9. The current density is given by the slope of the

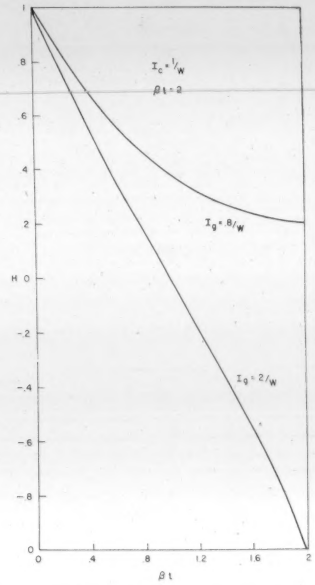


Fig. 9. Magnetic field variation through film for in-line cryotron gate.

curve. The small signal gain increases as βt increases as shown in Fig. 10 and approaches the critical field model for large βt. Actual control characteristics are shown in Fig. 11 and it is seen that they are somewhat similar to the theoretical predictions. The overall gain and the small signal gain are compared with the theory in Fig. 12. Similar characteristics are obtained if the control film is replaced by an external magnetic field.

The in-line cryotron introduces a resistance into the circuit that is proportional to the length, typical values being 0·8 ohm/inch. Therefore, the resistance can be made as large as desired at the expense of the space occupied and the energy dissipated. For this reason the in-line cryotron is

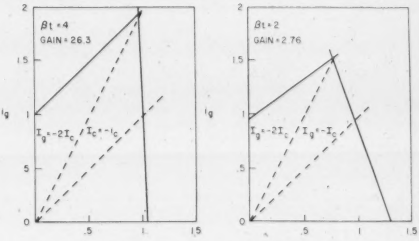


Fig. 10.—Effect of βt on gain characteristic.

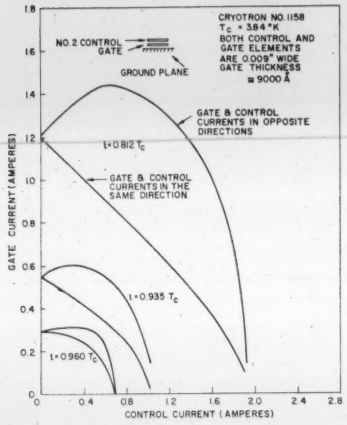

CRYOTRON NO. 1158
T_c = 3.84 °K
BOTH CONTROL AND GATE ELEMENTS ARE 0.009" WIDE
GATE THICKNESS ≈ 9000 Å
NO. 2 CONTROL GATE
GROUND PLANE
GATE & CONTROL CURRENTS IN OPPOSITE DIRECTIONS
GATE & CONTROL CURRENTS IN THE SAME DIRECTION
t = 0.812 T_c
t = 0.935 T_c
t = 0.960 T_c
GATE CURRENT (AMPERES)
CONTROL CURRENT (AMPERES)

Fig. 11. Actual characteristic of in-line cryotron.

particularly useful for driving long lines as would be encountered in a large memory array without an excessively large L/R time constant.

Future developments

The use of biasing techniques and the use of in-line cryotrons have significantly increased the operating speed of superconducting circuits. Simple loops have been constructed with transfers occurring in 10 ns (Rosenberger 1960). Current progress in alloy film technology (Toxen 1960) offers considerable promise for further improvements. The combination of these techniques should significantly increase the operating speed of cryogenic circuits in the future.

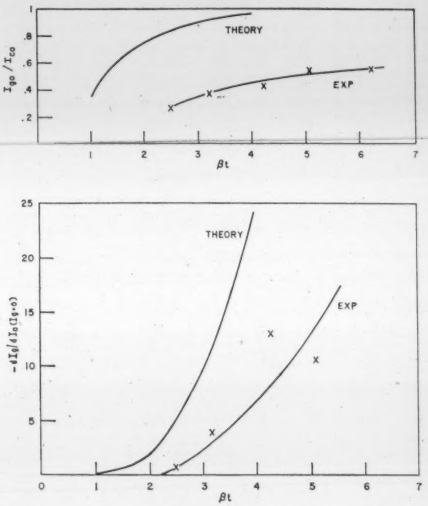


Fig. 12. Comparison of theory with experimental data for in-line cryotron.

References

BUCK, D. A., 1956, *Proc. Inst. Radio Engrs, N.Y.*, **44**, 482.
ROSENBERGER, G. B., 1960, *Solid State Electronics*, **1**, 388.
TOXEN, A. M., 1960, *Solid State Electronics*, **1**, 363.
YOUNG, D. R., 1959, *Progress in Cryogenics*, **1** (Heywood and Co. Ltd.), Chapter 1.
YOUNG, D. R., SWIHART, J. C., TANSAL, S., and MEYERS, N. H., 1960, *Solid State Electronics*, **1**, 378.

p–n junctions at very low temperatures

by A. K. JONSCHER, B.Sc., Ph.D., A.M.I.E.E., A.Inst.P., The General Electric Company Limited, Central Research Laboratories, Hirst Research Centre, Wembley, Middlesex

MS. received 23rd March 1961

Abstract

This paper gives a brief review of those low temperature properties of silicon and germanium which are relevant to the performance of p–n junctions at liquid helium temperatures. Experimental results are presented for germanium and silicon diodes and their interpretation is attempted. It is found that germanium devices behave in a manner similar to room temperature behaviour provided the ohmic contacts are adequate. Silicon diodes exhibit anomalies owing to insufficient doping of the end regions or to trapping in the base.

1. Introduction

THE subject of this paper is the electrical properties of p–n junctions at temperatures in the neighbourhood of 4° K and the discussion is limited to ordinary p–n junctions in silicon and germanium, to the exclusion of tunnel junctions. This limitation is justified by the fact that the performance of tunnel diodes changes very little between room temperature and 4° K and such phenomena as are observed have been the subject of recent reviews elsewhere (Hall 1960). The behaviour of conventional junctions, on the other hand, appears to have attracted much less attention so far, although some of the phenomena concerned are of both practical and theoretical interest.

This paper starts with a brief review of certain basic phenomena at very low temperatures, such as conduction of current, recombination, trapping and injection of excess carriers. It then presents some experimental results relating to germanium and silicon devices. An interpretation of these results is attempted.

2. Conduction of current in germanium and silicon at very low temperatures

The electrical conductivity σ in a semiconductor due to a density n_0 of free electrons may be expressed by the familiar formula

$$\sigma = e\mu_e n_0 \tag{1}$$

where e is the electronic charge and μ_e the electron mobility. At very low temperatures the magnitude of the conductivity ascribable to free carriers is therefore determined by the temperature dependence of n_0 and μ.

The density of free carriers in germanium at liquid helium temperature is very much lower than at room temperature but not necessarily negligible. In uncompensated n-type material containing a density N_D of donors not exceeding some 10^{15} cm^{-3} the equilibrium density of free electrons is given by (Smith 1959)

$$n_0 = \tfrac{1}{2}(N_D N_C)^{1/2} \exp(-\epsilon_d/2kT) \tag{2}$$

where ϵ_d is the activation energy of donors and

$$N_C = 2(2\pi m_e kT/h^2)^{3/2} \simeq 2 \times 10^{15} T^{3/2}$$

is the effective density of states. Eqn (2) is valid provided that any compensating minority impurity density N_A is less than n_0, otherwise the following relation applies:

$$n_0 = \frac{N_c(N_D - N_A)}{2N_A} \exp\left(-\frac{\epsilon_d}{kT}\right). \tag{3}$$

Carrier densities between 10^5 and 10^7 cm^{-3} would not be uncommon in uncompensated germanium at 4° K, considerably lower values applying to heavily compensated material.

The same formulae, when applied to silicon, give completely negligible carrier densities on account of the substantially smaller exponential factors.

With regard to mobilities of free carriers in silicon and germanium it is difficult to find experimental data at temperatures below 10° K. Some of the available data relating to higher temperatures are summarized in Fig. 1. The

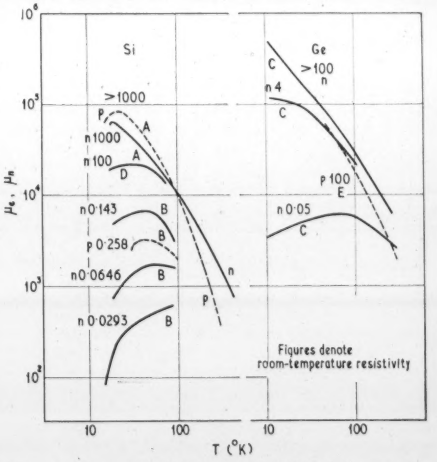


Fig. 1. Temperature dependence of mobility of electrons and holes in silicon and germanium.

A, Putley and Mitchell (1958); B, Swartz (1960); C, Debye and Conwell (1954); D, Logan and Peters (1960); E, Morin (1954).

importance of ionized impurity scattering which gives rise to a temperature dependence of the form

$$\mu \propto T^{3/2}$$

can be seen from these curves. Ionized impurity scattering is very sensitive to the presence of compensating minority

impurities, since a density N_A of such centres gives rise to a density $2N_A$ of oppositely charged ionized centres. It follows, therefore, that the value of room-temperature conductivity, being independent of the degree of compensation, gives no guidance as to the value of conductivity at liquid helium temperature.

The electrical conductivity may not necessarily follow the law given by Eqns (1)–(3) down to very low temperatures. When the free carrier density becomes very low due to the presence of compensation or to a high impurity activation energy, a different conduction mechanism may become dominant. This is due to 'hopping' of bound electrons from occupied impurity centres to neighbouring unoccupied ones without the need for excitation into the free band. This mechanism depends very critically upon the density of majority impurity and upon the degree of compensation (Miller and Abrahams 1960, Fritzsche and Cuevas 1960). Its temperature dependence may be expressed by the formula

$$\sigma \propto \exp(-\epsilon_3/kT)$$

which is similar to Eqns (2) and (3), but the 'activation energy' ϵ_3 in compensated material is much lower than the impurity activation energy ϵ_d. This mechanism effectively sets a lower limit to the conductivity at low temperatures in germanium and silicon.

At sufficiently high impurity densities, of the order 10^{16} cm^{-3} in germanium and 10^{18} cm^{-3} in silicon, the mutual interaction of impurity levels gives rise to the formation of a quasi-continuous band of energy levels. At the same time, the shape of the lower edge of the conduction band may be

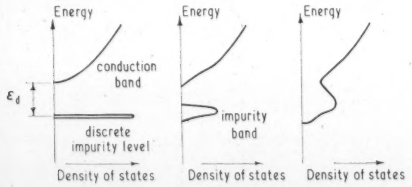


Fig. 2. Impurity band formation (diagrammatic) as a function of the density of the majority impurity.
(*a*) low; (*b*) intermediate; (*c*) high.

substantially affected (Fig. 2). One consequence of this state of affairs is a reduction of the impurity activation energy. The impurity band, so long as it remains separate from the conduction band, should be completely filled in the absence of compensating impurities. The presence of such impurities would lead to a decrease of electron density in the impurity band. However, experimental results for n-type germanium (Fritzsche 1958) suggest that addition of compensating impurities *reduces* the impurity band conductivity, which is not compatible with the simple model outlined above. Measurements on n- and p-type silicon (Swartz 1960) are not conclusive because the degree of compensation was unknown.

Further increase of impurity density, to 10^{18} cm^{-3} in germanium and in excess of 10^{19} cm^{-3} in silicon, leads finally to complete merging of the conduction and impurity bands (Fig. 2(*c*)) and the formation of a fully degenerate, quasi-metallic electron population. This state of affairs has as a consequence an effective reduction of the energy gap between the (non-degenerate) minority band and the degenerate majority band.

To sum up, the following points relating to the electrical properties of silicon and germanium at very low temperatures are relevant to the present discussion of p–n junctions. Provided the material is non-degenerate, the conductivity is several orders of magnitude less than at room temperature. The precise value of conductivity at a given low temperature depends rather critically on the compensation and may vary over several orders of magnitude between different specimens of the same room temperature conductivity. The mobility of majority carriers, in so far as it corresponds to motion in the conduction band, varies over wide limits and is also very sensitive to the extent of compensation. The mobility in the impurity band may be much lower than in the conduction band.

An important new phenomenon sets in at electric fields exceeding a certain critical value—in germanium this may lie between 5 and 150 v/cm, depending on compensation and doping level. Above this critical field the stray free carriers acquire sufficient energy to ionize the impurities by impact and cause an avalanche-like increase of conductivity, shown in Fig. 3, to values approaching the room-temperature order (Koenig 1958, 1959, McWhorter and Rediker 1959). This effect is especially important in germanium, although it is reported to have been observed in silicon at much higher fields—10^2–10^3 v/cm (Kaiser and Wheatley 1959).

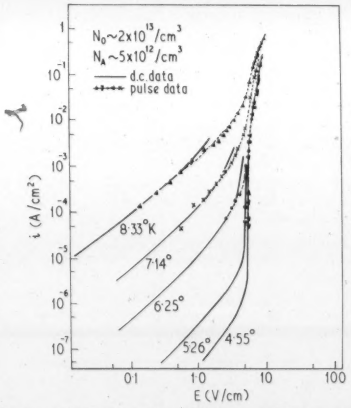


Fig. 3. Conduction of current in n-type germanium (*after* S. H. Koenig 1958).

The dynamic equilibrium between the rate of ionization by impact and the rate of recombination by capture of free carriers in ionized impurities determines the degree of ionization of impurities as a function of the electric field.

3. Generation–recombination relations

The equilibrium densities n_0 and p_0 of free carriers in the conduction and the valence bands are maintained by a two-way traffic of electrons between these bands. Any attempt to upset these equilibrium densities by withdrawal of either or both carrier species or by injection of excess carriers is

immediately counteracted by a net generation or recombination, respectively.

Interband transitions may be direct or by way of intermediate levels, referred to as recombination centres. The latter transitions are by far the more probable and they may be described in terms of the familiar relation (Rose 1957)

$$dt/\tau = v\sigma N dt$$

where dt/τ is the probability that a carrier will be recombined on one of the recombination centres in a time interval dt, v is the average thermal velocity of the carrier, σ is the capture cross section of the recombination centres for the relevant carrier species and N is the number of centres capable of capturing the carrier. The quantity τ is frequently referred to as lifetime.

The actual process of recombination on one intermediate level is a two-stage process: thus a free electron may be captured first and the centre becomes recharged, whereupon a hole is captured and the centre returns to its initial state. Two capture cross sections are, therefore, involved and they are likely to be very different since in one of them the centre is usually neutral, while in the other it is attractive to the other carrier. Thus we may have reactions of the type

$$\begin{aligned} \text{electron} + C^{\times} &\rightarrow C^{-} \\ C^{-} + \text{hole} &\rightarrow C^{\times} \end{aligned}$$

where $C^{\times}$ and C^{-} refer, respectively, to the neutral and negatively charged states of the capturing centre.

The total rate at which electrons are being captured is $nv_e\sigma_e N_e$, while the total rate for holes is $pv_h\sigma_h N_h$. In steady state both rates must be equal, so that:

$$nv_e\sigma_e N_e = pv_h\sigma_h N_h$$

with the obvious additional condition:

$$N_e + N_h = N_{\text{total}}.$$

Knowing that these recombination–generation rates lead to values of lifetime in the range 10^{-7}–10^{-3} s in silicon and germanium at ordinary temperatures, the question arises of the likely values of lifetime at very low temperatures. p and n are, of course, very strongly temperature dependent and the velocity v is proportional to $T^{1/2}$. The capture cross sections σ_h and σ_e may in some cases decrease very rapidly with temperature (Lax 1960) but, on the scant available evidence (Lax 1960, Bemski 1958), may in other cases remain sensibly temperature independent (Fig. 4). For example, while the capture cross sections Au_e^{+} and Au_h^{-} increase rapidly as the temperature falls, the cross sections $Au_h^{\times}$ and $Au_e^{\times}$ for the complementary transitions retain their order of magnitude. The latter will therefore remain the rate-determining quantities in the recombination process. Thus, on balance, bearing in mind the reduction of carrier densities, one concludes that the thermal transition rates in the bulk are very much reduced at low temperatures compared with room temperature. This result is, of course, compatible with the scarcity of energetic phonons which could provide the energy for these transitions. A similar conclusion applies to the surface transition rates.

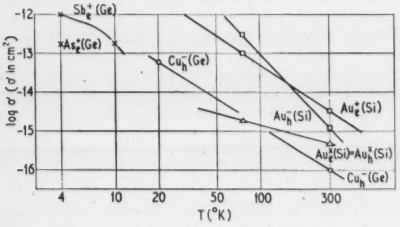


Fig. 4. Trapping capture cross sections in silicon and germanium.

One would therefore expect that all device phenomena depending upon thermal *generation* due to the removal of one or both carrier species, whether at the surface or in the bulk, would be very appreciably slowed down. Examples of these are, in particular, space charge region generation in reverse-biased junctions and the operation of so-called 'ohmic' metal-semiconductor contacts.

It is otherwise with *recombination* in the presence of injected excess carrier densities. Here p and n may well reach values comparable with those obtainable under room-temperature operation. Thus the recombination rates for equal injected carrier densities may be of similar orders of magnitude at room temperature and at liquid helium temperature.

There is no experimental evidence concerning lifetime at liquid helium temperature. Broad trends discernible in the range 100–300° K indicate a decrease of lifetime with temperature (Kalashnikov 1959, Goldstein, Mette and Gärtner 1959) or, as with Au recombination centres, an almost constant value. Experimental evidence relating to the performance of diodes and transistors at liquid helium temperature suggests that lifetime in germanium at reasonable injection levels is not orders of magnitude lower than at room temperature.

The very large increase of some capture cross sections, especially those for shallow impurity levels, whilst not necessarily entailing a comparable increase of the recombination rate, nevertheless implies heavy trapping of carriers in these states. At the same time, the low generation rate means low probability of re-excitation from these traps—once trapped, carriers may remain in deep traps for hours or as long as the temperature is kept low.

Trapping on ionized donor or acceptor levels may become especially pronounced at high injection levels, thus countering the ionizing effect of the electric field in germanium.

4. Metal–semiconductor contacts

At normal temperature certain metal–semiconductor contacts may be presumed to possess 'ohmic' properties. This implies that they can supply majority carriers at any required rate with negligible potential drop across the contact. These requirements can be satisfied if the contact has a very high generation–recombination rate and if any potential barrier associated with it is small in comparison with the thermal energy kT. The latter condition is usually met in low-resistivity material and hence the emphasis on n–n_+–metal and p–p_+–metal structures as ohmic contacts.

The situation can be radically altered at very low temperatures by the fact that, on the one hand, the thermal generation rate goes down very appreciably and, on the other hand, the importance of even small barriers is greatly magnified by the smallness of kT.

The application of ohmic contacts for low temperature operation to germanium does not present any serious difficulties. All alloyed contacts and some soldered ones are adequate for most purposes. This is at least partly due to

the small activation energy of impurities in this material. In silicon, on the other hand, with four times the activation energy of germanium, the situation is rather different. Ordinary soldered contacts are seldom adequate. More surprisingly, perhaps, even alloyed contacts appear in many cases to be non-ohmic and the safe way of obtaining a good ohmic contact at very low temperatures between non-degenerate silicon and a metal is by the intermediary of a strongly degenerate layer of the same type of conductivity. This implies an impurity density in excess of 10^{19} cm^{-3}.

5. Injection of excess carrier densities

The injection of excess carrier densities at very low temperatures follows the same rules (Jonscher 1960) as at room temperature, but some of these rules tend to be sufficiently taken for granted to be overlooked. The essential criteria for injection are:

(i) The presence of a source of free carriers—this implies a fully degenerate region at least on one side of the junction.

(ii) The reduction of the diffusion potential to a value not exceeding several kT—that means some millivolts.

(iii) The availability of excess majority carriers from the 'back' contact, to neutralize first the fixed space charge and then the injected minority carriers.

This requires an 'ohmic' contact at the other end.

Once injected the excess carriers diffuse, drift and decay as usual but certain complications may arise. The drift mobility may differ from the relevant mobility in the absence of injection if the latter corresponds to impurity band conduction, since minority carriers are *not* subject to impurity band effects. Likewise, it is possible that at sufficiently high injection levels some of the compensating majority charge may no longer be confined to the impurity band and thus may contribute to a higher effective mobility. The diffusion constant would be expected to obey Einstein's relation:

$$D = \mu kT/e$$

provided the injected carriers are non-degenerate. This assumption, of course, may well be violated even at injection levels which are moderately low by normal standards. The mobility μ in this expression is that appropriate to the minority carriers. The low-temperature value of the diffusion constant may be quite low regardless of the potentially high values of the minority mobility on account of the smallness of kT.

Recently published experimental and theoretical work (Paige 1960, McLean and Paige 1960) relating to the drift mobility of minority carriers in germanium at very low temperatures reveals the importance of electron–hole scattering on the transport of minority carriers. For majority electrons or holes, carrier–carrier scattering only gives rise to a redistribution of momentum between them and therefore has a small effect on mobility. On the other hand, injected minority carriers that are scattered by the majority carriers experience a more drastic reduction of mobility. The latter may be dominated by electron–hole scattering and it is predicted that under favourable circumstances, e.g. in InSb, this drag effect may lead to negative mobilities for injected minority holes.

6. Voltage-current characteristics of germanium p–n junctions

Figure 5 gives the forward voltage–current characteristics of alloyed germanium diodes at 300, 77 and 4·2° K. The points to note regarding the behaviour at 77° K are the logarithmic slope lying between e/kT and $e/2kT$ and the presence of cusps, especially on the characteristic of No. 1 diode. Both points may be connected with the fact that the dominant mechanism at low currents is most likely recombination in the space charge region.

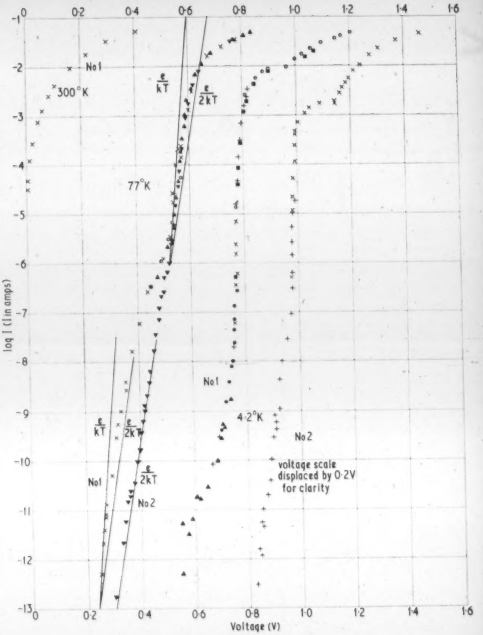


Fig. 5. Voltage–current characteristics of alloyed germanium diodes at 300, 77 and 4·2° K.

The characteristics at 4·2° K show a comparatively slow exponential rise up to about 0·75 V (i.e. almost the energy gap). If this were interpreted in terms of the usual diode law $I \propto \exp(eV/kT)$, the relevant temperature deduced from the slope would be in the range 150–300° K. The slope becomes much steeper at 10^{-7} ampere, but even then the characteristic temperature is as high as 30–40° K. These facts clearly imply that the conduction can be due neither to ordinary injection nor to space charge region recombination as it is known under normal conditions (Jonscher 1960, § 4.6).

The rate of rise of current eventually decreases, but a distinct cusp can be discerned on the characteristics at 0·95 V leading to further increase of current.

The interpretation of these characteristics is indicated in Fig. 6. At very low currents there is no injection and some form of recombination dominates the behaviour. The majority current causes a field in the base region which is too low to give rise to impurity ionization avalanche. At some stage on the steeply rising branch of the characteristic, depending upon the doping level and compensation, the majority current becomes sufficiently high for avalanche

multiplication to set in. This, however, has little effect on the characteristic, since it results only in the stabilization of the voltage drop across the base region.

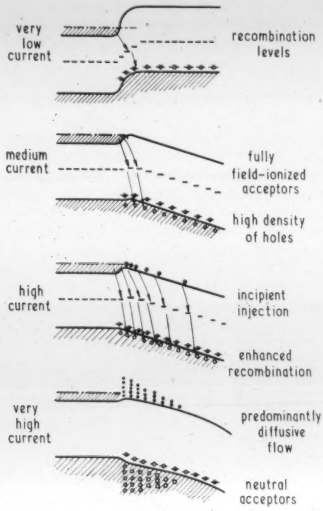


Fig. 6. Forward current flow in a germanium p–n junction.

Once the potential barrier has been substantially flattened the current due to the recombination in the space charge region will tend to saturate and at the same time injection will set in. Recombination of the minority carriers may now extend into the base region but the injection level is still low and diffusion in the base region is not yet significant.

It is evident, however, that this situation cannot be maintained as the bias is further increased. At some stage injection into the base region must become sufficiently copious to cause a substantial increase of the diffusion component of the minority current and equally a substantial increase of the compensating majority density above the fully ionized density. This has two effects: firstly, conductivity modulation becomes appreciable and the current may now increase again beyond the value corresponding to unmodulated flow. Secondly, the presence of a high compensating density of majority carriers, especially near the junction, causes gradual de-ionization of the field ionized acceptors, leading to a reduction of the effectiveness of ionized impurity scattering. These are thought to be the reasons for the appearance of the distinct cusps on the upper parts of the characteristics. That part is plotted as $I^{\frac{1}{2}}$ against V in Fig. 7 and shows a reasonably straight line portion indicating the predominance of diffusive transport with removal at the end contact. This is in agreement with similar results obtained with the same diodes at other temperatures and shown in Figs 7 and 8. A relation of the type $I \propto (V - V_0)^2$ was found to be obeyed quite generally by all diodes in which the current flow was dominated by injection and removal of carriers at the back contact rather than by bulk recombination (Jonscher 1958).

It is interesting to contrast the behaviour of a p–n junction diode described above with that of a structure in which the supply of majority carriers is restricted by a series resistance. An ordinary junction transistor was used as a diode, either

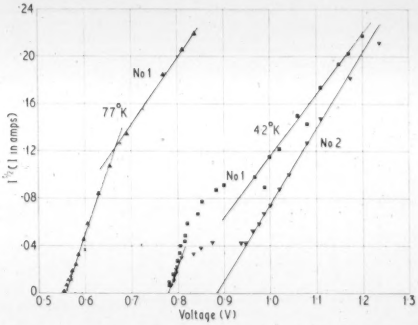


Fig. 7. $I^{\frac{1}{2}}$ plotted against V for alloyed germanium diode at 77 and 4·2° K.

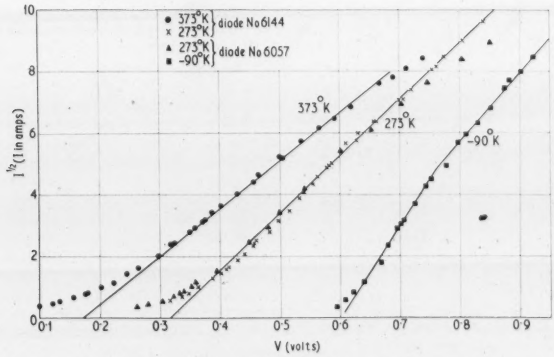


Fig. 8. Forward characteristics (I, V) of alloyed germanium diodes (*after* Jonscher 1958).

with base and collector strapped together and emitter used as the junction, or with emitter and base strapped together and collector used as the junction. At room temperature this structure would be indistinguishable from a diode, the forward current flow being due mainly to injection of carriers from the forward biased junction and their collection at the short-circuited junction. At helium temperature, however, the flow is due primarily to the majority carriers which can be supplied only by the peripheral base terminal and not by the short-circuited p–n junction. The characteristic shown in Fig. 9 is, therefore, determined at first by the ohmic resistance in the material between the base and either the collector or the emitter, as the case may be. At a certain voltage, depending upon the length of this path, impurity avalanche breakdown sets in and the current rises very steeply with slight tendency to negative resistance. The latter may be due to slight compensation in the material of the base region (McWhorter and Rediker 1959).

The reverse characteristics of germanium diodes show breakdown voltages of the same order of magnitude as at room-temperature. This suggests that the prevailing breakdown mechanism, presumably avalanche breakdown, remains largely unaffected by the lowering of the temperature. The reverse current, while very much lower than at room temperature is difficult to measure, however, since very large transients interfere with measurements. This may be due to ionization of discrete trapping levels leading to localized avalanches.

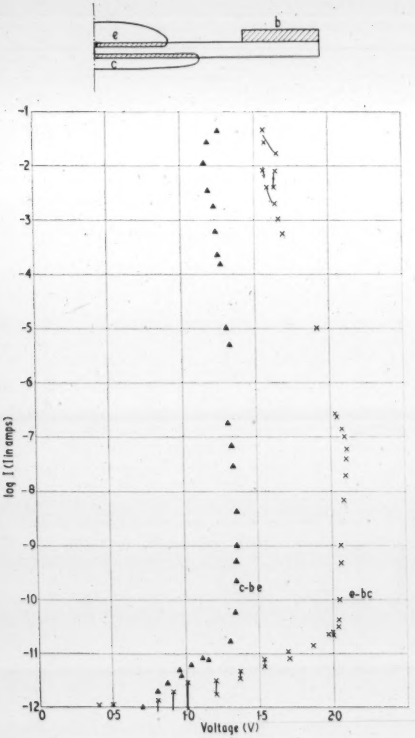


Fig. 9. Voltage–current characteristics of a germanium alloyed transistor used as a diode.

7. Characteristics of silicon diodes

We have studied a large variety of commercially available silicon devices, both alloyed and diffused. Their behaviour was quite unlike that of germanium devices and two distinct patterns could be discerned. All *alloyed* devices showed more or less 'back-to-back' characteristics with clear breakdown at an appreciable voltage in either direction. On the reverse branch this voltage would bear some relation to the room-temperature breakdown voltage; on what would have been the forward branch breakdown voltages of 20–100 volts would be typical with little or no tendency to negative resistance. In all cases the sustaining voltage would be in excess of 20 volts. This behaviour might be compared with that of the germanium 'transistor diode' shown in Fig. 9. The cause of the high sustaining voltage is most likely the failure to inject minority carriers into the base region, because of insufficient doping in the recrystallized region either on the junction side or in the back contact, or both. This explanation seems plausible in view of the fact that to achieve full degeneracy in silicon impurity densities in excess of 10^{19} cm^{-3} may be necessary.

An entirely different behaviour was observed in some types of *diffused* diodes and is shown schematically in Fig. 10

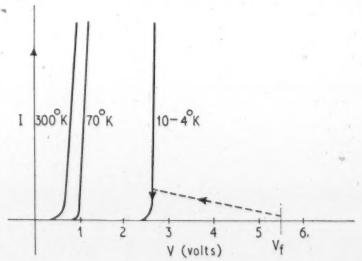


Fig. 10. Forward characteristics of diffused silicon junction diodes at different temperatures with 50 c/s display.

which relates to a typical 50 c/s a.c. display. As the temperature is lowered towards 4° K a distinct peak appears on the forward branch, followed by breakdown to a sustaining voltage of between 2 and 3 volts. On the return path there is much less, if any, tendency to switching. The magnitude of the peak depends on the voltage amplitude of the reverse half-cycle: the more the reverse voltage exceeds a critical threshold value the higher the peak.

This effect was studied under d.c. conditions and it was found that on first immersion from a higher temperature the forward current would be less than 10^{-11} ampere up to approximately one volt of a well-defined breakdown voltage. This would be reproducible for each diode and would lie between 8 and 40 volts typically. After breakdown and removal of bias, subsequent re-application of forward bias would give only a small peak, from 2·5 to 6 volts, according to the magnitude of the original peak but independently of the duration of time in the off state. The original switching

could only be restored by momentarily raising the temperature above, say, 20° K or by applying a reverse bias in excess of a certain minimum value, say 150 volts.

In the experiment the current was measured by observing the potential across a resistance in series with the diode. Other measurements were taken with this resistance replaced by a capacitor. The pre-breakdown currents in the forward and reverse directions are so small that it is possible to apply moderate biases indefinitely without collecting any charge on the condenser other than that due to the displacement current in the system. On removal of bias there would be no residual charge left on the condenser. If, however, the device were first broken down in the forward direction and then biased in the reverse direction with the condenser in series, one

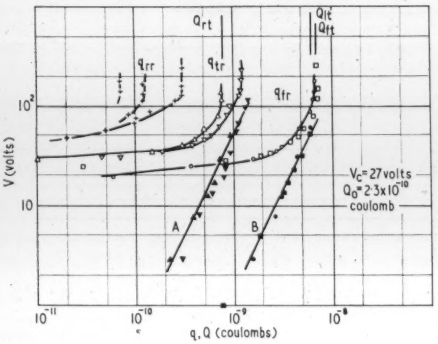


Fig. 11. Reverse charge transfer measurements for diode No. 48.

Curves A, B: $q_{tr} + Q_0$ and q_{fr} respectively plotted against $V - V_c$.

would obtain an irreversible withdrawal of charge as shown by the open squares on Fig. 11. The symbol q denotes charge withdrawn irreversibly as a function of the applied bias. The two subscripts denote, respectively, the immediate prior state and the actual state of the device, with the following abbreviations: t, thermal equilibrium; f, forward breakdown; r, reverse bias; l, illumination with light. A threshold reverse bias of some 30 volts would be required, beyond which the amount of extracted charge q_{fr} would increase rapidly but would eventually saturate. If the forward bias were now applied the breakdown voltage V_{rf} would be found to be considerably in excess of the original breakdown voltages perhaps in the range 12–80 volts.

Returning to charge withdrawal, however, if the reverse bias were applied not following breakdown but immediately after immersion from thermal equilibrium, the charge withdrawn, q_{tr}, would be as given by the open triangles on the graph. The saturation charge would be markedly lower and the threshold bias higher than after forward breakdown. Finally, if an attempt were made to withdraw charge again after previous reverse biasing to saturation, q_{rr}, the curves marked with crosses would be obtained.

It was found that illumination with light had exactly the same effect as breakdown in the forward direction.

It was possible to obtain recovery of charge in unbiased condition by warming up the device. The charge so obtained is denoted by Q with appropriate subscripts. Thus Q_{ft} is the charge recovered by warming up, i.e. restoring thermal equilibrium after forward breakdown, while Q_{lt} is the charge recovered after previous illumination. It is clear from the graph that these two are equal and are the same as the saturation charge q_{fr} withdrawn on reverse biasing. Likewise the charge Q_{rt} obtained on warming up a device previously reverse-biased turns out to be the same as the saturation value of q_{tr} recovered on reverse-biasing from thermal equilibrium.

It is also possible to measure a definite transfer of charge in the forward direction prior to breakdown, Fig. 12. This in turn depends on whether the forward bias follows reverse bias, when more charge is required to cause breakdown, q_{rf}, or whether it follows thermal equilibrium conditions when less charge is necessary, q_{tf}.

To sum up, depending upon their past history, the diodes in question can exist in different states of charging, the equilibrium state, post-breakdown state and post-reverse-bias state. The forward breakdown voltages V_{tf}, V_{ff} and V_{rf}, respectively, are different according to the state. The device possesses memory and remains in each state until a transfer of charge is induced by suitable biasing or illumination.

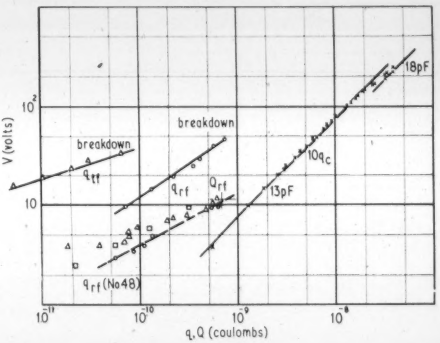


Fig. 12. Forward charge transfer and capacitive charge in diode No. 186.

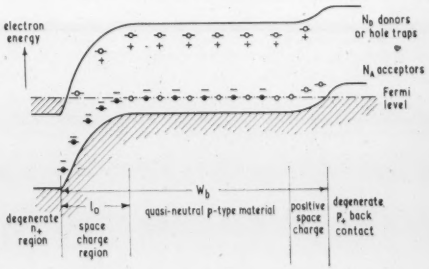


Fig. 13. p–n junction at very low temperature in thermal equilibrium.

An explanation of this behaviour is offered in terms of a model involving compensated material in the base region of the diode, as shown in Fig. 13. The outer p- and n-regions are fully degenerate. Potential barriers which are substantial

for the temperature in question, exist at either end—for the minority electrons at the junction and for the majority holes at the 'ohmic' contact. Thus, initially, neither carrier may enter the base region.

The application of a 'forward' bias to this structure results in the situation represented in Fig. 14(*a*). Any holes entering from the p-region on the right are immediately trapped on ionized acceptors and their space charge helps to build up the barrier to further flow. At any particular bias only a shift of charge occurs by 'hopping' in the incompletely filled

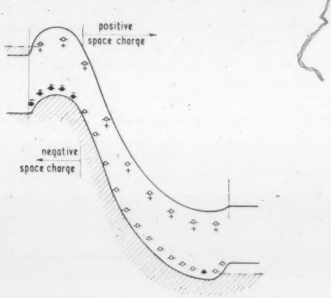


Fig. 14(*a*). The pre-breakdown condition under forward bias.

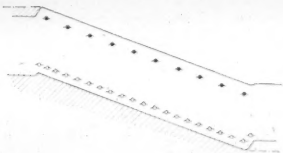


Fig. 14(*b*). Post-breakdown condition under forward bias. Large injected densities of both carrier species maintain 'inverted' impurity level populations.

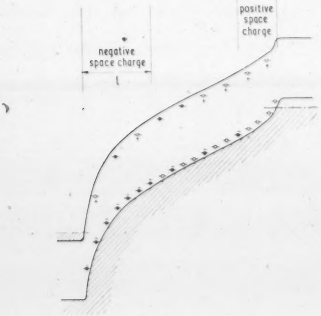


Fig. 14(*c*). Reverse biased condition.

majority acceptor levels, the density of which is of the order of 10^{14} cm^{-3}, but continuous conduction is not possible. Finally, of course, a critical voltage is reached at which the small barrier on the right is completely eliminated and a copious flow of holes ensues. These holes neutralize the junction space charge on the left and open the gate to unrestricted flow of electrons. This corresponds to a collapse of voltage across the structure and the final situation is as shown in Fig. 14(*b*). This is very similar to current flow in a p–i–n diode. It should be noted that the donors and acceptors are now completely neutral instead of mostly ionized, as in thermal equilibrium. This is the charge difference between the two states and it appears as the charge Q_{ft} recovered on raising the temperature.

It is evident that the same condition as shown in this figure can be brought about by illumination.

Once neutralized, the donors and acceptors cannot revert to their original state even though the current may be interrupted; the trapped carriers will remain so indefinitely and the forward characteristic will show no breakdown.

A return to the normal condition is possible by heating the device, thus supplying the thermal energy to release the trapped charge. Alternatively, the application of the reverse bias can have the same effect, as shown in Fig. 14(*c*). It is believed that field ionization of impurities is responsible for the transfer of charge in this case. This accounts for the existence of a critical reverse threshold voltage for charge transfer. The corresponding threshold field is of the order of a few kilovolts per cm. It is easy to see that the application of a high reverse bias leads to an increase of the forward breakdown voltage above the value corresponding to thermal equilibrium conditions.

Detailed analysis was carried out for a model in which the densities of donors and acceptors are N_{D} and N_{A} respectively and charge recovery is assumed to set in at a critical field E_{c}. It follows then that the charge q_{ft} recovered on reverse biasing after forward bias is given by the expression

$$q_{\mathrm{ft}} = Ae\left(\frac{\kappa}{2\pi}\right)^{1/2} \frac{N_{\mathrm{A}} + N_{\mathrm{D}}}{(N_{\mathrm{A}} - N_{\mathrm{D}})^{1/2}}(V - V_{\mathrm{c}})^{1/2}$$

where $V_{\mathrm{c}} = E_{\mathrm{c}}W_{\mathrm{b}}$ is the threshold voltage, W_{b} is the width of the base region, κ is the dielectric constant and A is the area of the junction. The charge q_{tr} recovered from thermal equilibrium is similarly given by

$$q_{\mathrm{tr}} + Q_0 = Ae\left(\frac{\kappa}{2\pi}\right)^{1/2} (N_{\mathrm{A}} - N_{\mathrm{D}})^{1/2}(V - V_{\mathrm{c}})^{1/2}$$

where Q_0 is the charge present under thermal equilibrium conditions in the space charge region of the junction.

The solid points in Figs 11 and 12 represent graphs of q_{ft} plotted against $V - V_{\mathrm{c}}$ and $q_{\mathrm{tr}} + Q_0$ plotted against $V - V_{\mathrm{c}}$, where V_{c} and Q_0 have been suitably chosen to give a slope approximating to $\frac{1}{2}$. The fit of points to a straight line throughout the range appears to support the simple model chosen for interpreting the results.

8. Capacitance of p–n junctions

As might be expected, the capacitance of the space charge region is unaffected by the lowering of the temperature, since all impurity centres remain fully ionized. The capacitance of the diode as a whole, however, suffers a reduction by two or more orders of magnitude as soon as the conductivity of the bulk regions on at least one side of the junction falls to a sufficiently low value for the capacitance of this region to become important (Vul and Zavaritskaya 1960). This happens below 10° K in germanium and below 30° K in silicon. The low temperature capacitance corresponds to the total thickness of material between metal

contacts or between degenerate regions. No change with temperature would be expected, of course, if both regions were degenerate, as in a tunnel diode.

References

BEMSKI, G., 1958, *Proc. Inst. Radio Engrs, N.Y.*, **46,** 990.

DEBYE, P. P., and CONWELL, E. M., 1954, *Phys. Rev.*, **93,** 693.

FRITZSCHE, H., 1958, *J. Phys. Chem. Solids*, **6,** 69.

FRITZSCHE, H., and CUEVAS, M., 1960, *Phys. Rev.*, **119,** 1238.

GOLDSTEIN, S., METTE, H., and GÄRTNER, W. W., 1959, *J. Phys. Chem. Solids*, **8,** 78.

HALL, R. N., 1960, *Proceedings of the Prague Conference* (Prague: Czechoslovak Academy of Sciences).

JONSCHER, A. K., 1958, *J. Electronics and Control*, **5,** 226.

—— 1960, *Principles of Semiconductor Device Operation* (London: G. Bell and Sons).

KAISER, W., and WHEATLEY, G. H., 1959, *Phys. Rev. Letters*, **3,** 334.

KALASHNIKOV, S. G., 1959, *J. Phys. Chem. Solids*, **8,** 52.

KOENIG, S. H., 1958, *Phys. Rev.*, **110,** 986.

—— 1959, *J. Phys. Chem. Solids*, **8,** 227.

LAX, M., 1960, *Phys. Rev.*, **119,** 1502.

LOGAN, R. A., and PETERS, A. J., 1960, *J. Appl. Phys.*, **31,** 122.

MCWHORTER, A. L., and REDIKER, R. H., 1959, *Proc. Inst. Radio Engrs, N.Y.*, **47,** 1207.

MILLER, A., and ABRAHAMS, E., 1960, *Phys. Rev.*, **120,** 745.

MORIN, F. J., 1954, *Phys. Rev.*, **93,** 62.

PUTLEY, E. H., and MITCHELL, W. H., 1958, *Proc. Phys. Soc.*, **72,** 193.

ROSE, A., 1957, *Progress in Semiconductors*, Vol. 2 (London: Heywood), p. 111.

SMITH, R. A., 1959, *Semiconductors* (Cambridge: University Press), p. 90.

SWARTZ, G. A., 1960, *J. Phys. Chem. Solids*, **12,** 245.

VUL, B. M., and ZAVARITSKAYA, E. I., 1960, *Soviet Physics T.E.T.P.*, **11,** 6.

The electricity of precipitation

by J. ALAN CHALMERS, M.A., Ph.D., F.Inst.P., Department of Physics, Durham Colleges in the University of Durham, South Road, Durham City

MS. received 19*th May* 1961

Abstract

Results of measurements on precipitation electricity are described and the problems of explaining them are discussed.

1. Introduction

THERE are a number of points of interest in the electrification of precipitation, including rain, hail and snow. In the first place, the charge on precipitation must be taken into account in considering the total charge reaching the earth's surface. Then, it may well be that the charge on the precipitation falling out of a cloud can give useful indications as to the charge-separation processes at work in the cloud. And relations between the precipitation current and other factors can yield information as to how the precipitation must have obtained its charge, and hence what charging processes are operative.

Precipitation, both rain and snow, may be broadly divided into: (*a*) stormy and showery precipitation mainly from cumulo-nimbus clouds, and (*b*) continuous precipitation mainly from nimbo-stratus clouds. It is found that this division is suitable for the discussion of electrical effects, mainly because point discharge occurs usually during stormy conditions and seldom during continuous precipitation; these two conditions will be described as 'stormy' and 'quiet'.

2. Methods of measurement

It is not proposed here to give any details of the measurement techniques, for which reference may be made to original papers, but rather to distinguish between the different types of method.

Measurements may be made either of the charges on individual raindrops, snowflakes or hailstones or of the total current reaching an area and comprising a number of single drops; and this current can be measured either as an actual current or as the charge collected over a certain period of time, giving the mean current.

When the current is measured, there is the choice of using an open receiver or a shielded receiver. An open receiver can be made so as to simulate a portion of the earth's surface and give conditions close to the natural, but there is the complication that, in addition to the true precipitation current, there are measured also any ionic conduction currents and effects due to changes of potential gradient; these last can be understood if we think of lines of force in the atmosphere ending on bound charges on the surface of the receiver; then, if there is a change in the potential gradient the bound charge alters and there is a current through the measuring apparatus which is not distinguishable from a true current. Because of this difficulty, a shielded receiver has often been used, such that the lines of force end on the earthed shield rather than on the receiver itself. However, a method of compensating for the bound-charge current was developed by Adamson (1960) using a field machine to give the effect of the potential gradient alone and oppose the effect of this on the current; this method was used by Ramsay and Chalmers (1960).

In earlier work (Elster and Geitel 1888 and many others), a quadrant or other electrometer was used for measurement; later, a sensitive galvanometer was used (Herath 1914, Chalmers and Little 1947) and more recently electronic methods have been widely used (Ramsay and Chalmers 1960, etc.).

For the measurement of charges on single drops, earlier work used a collector connected to a sensitive electrometer (Gschwend 1922); later, electronic methods were used (Chalmers and Pasquill 1938, etc.). Still more recently, the technique has been to allow the drops to fall through a ring, often connected to a cathode-ray oscilloscope (Gunn 1949, etc.); in this way it can be ensured that there is no effect of splashing, etc., of the drop.

3. Auxiliary measurements

In addition to the actual measurement of the precipitation current, various other quantities have been measured. In many cases the potential gradient has been measured and also the point-discharge current when the potential gradient was sufficiently high. Many workers have measured the rate of rainfall; in the case of a shielded receiver, it has been possible to measure the actual amount of water reaching the receiver (e.g. McClelland and Nolan 1912); in other cases, particularly with an exposed receiver, independent apparatus has been used to measure the rate of rainfall.

For single drops, the size of each drop has been measured in various ways. Some (e.g. Gschwend 1922) have measured the size of the stain produced by each drop on prepared paper. Others (e.g. Gunn 1949) have found the time for the drop to fall between two induction rings. For larger drops, Smith (1955) used the variation in the capacity of a parallel-plate condenser as a drop fell through. Arabadji (1959) measured the impulse of the fall of the drop on a piezo-electric crystal.

4. General results

One of the most obvious points of interest in precipitation electricity is that of the balance of charge brought to the ground by precipitation. This is often expressed in terms of the ratio of positive to negative charge received, but it can easily be seen that this ratio has a simple meaning only when the charges on single drops are considered; in any other case, each measurement probably includes drops of both signs. The results for this ratio have usually shown an

excess of positive charge coming down in most types of precipitation, but there are some observations of negative excess charge in stormy weather, more particularly in the cases of showers without lightning. Snow is also often found to show a negative excess of charge, especially in quiet conditions. For quiet continuous rain, the positive excess is very strongly marked.

The actual vertical current brought down to a given area of the ground by precipitation has been determined and gives results varying from 10^{-12} A/m^2 for light quiet rain up to as much as 10^{-8} A/m^2 in stormy conditions.

Results for the charges per unit volume of water vary from 10^{-6} C/m^3 to 10^{-2} C/m^3. The charges on single drops vary from less than 10^{-15} C up to over 10^{-10} C.

One of the most striking results of the measurement of single-drop charges is the wide divergence of values obtained even for drops of the same size arriving at the same time (Smith 1955). Nevertheless, two receivers placed close to one another and receiving rain currents over areas of the order of 10^{-2}m^2 show very nearly the same currents, so that the differences in the charges on single drops must be statistical.

5. Comparison of precipitation with other currents

During stormy weather, it is seldom that the precipitation current reaches as much as 10^{-8} A/m^2 and for the whole area of a storm the average is probably less than 10^{-9} A/m^2. The area covered by a storm is between 10 and 100 km^2, so that the precipitation current in the whole storm is certainly less than 10^{-1} A and possibly as little as 10^{-2} A. It is known, from observations above clouds (Gish and Wait 1950, Stergis, Rein and Kangas 1957) that the total current through a storm cloud is of the order of 1 A, so that the precipitation current can be only a small fraction of the total. This is rather surprising in view of the theories which postulate precipitation as an important agency in the separation of charge in clouds. However, it has been clearly shown (e.g. Schonland 1928) that point-discharge currents are the most important agency in the transfer of charge from clouds to the earth, while lightning comes next; not only are the effects of rain currents small but they are actually of the opposite sign to the main transfer. If precipitation is concerned in bringing down, inside the cloud, the negative charge which reaches the base of the cloud, there must be some process below the cloud by which the precipitation charges are reduced and often altered in sign.

During quiet precipitation, when there is no lightning and usually no point discharge, the relative importance of the precipitation current must be greater. In such cases, the precipitation current is more likely to give an indication of the charge separation current within the cloud, than in the cases of storms.

Estimates have been made of the total current brought down to 1 km^2 of the earth by precipitation; Wormell (1930) estimated 20 C per year for Cambridge. For an average for the whole world, Wait (1950) made a similar estimate and Israël (1953) suggested 30 C, but these can only be tentative guesses, since there is insufficient knowledge of precipitation currents all over the world.

6. Relation with point-discharge current in storms

Simpson (1949) found that, when conditions in stormy weather are fairly steady, the rain current is proportional to the point-discharge current but of opposite sign. Simpson found that the ratio of the two currents increases with the rate of rainfall, but tends towards a limiting value for high rates of rainfall.

This can be explained as follows: suppose there is a negative potential gradient so that point discharge occurs and brings a negative charge to the earth; the corresponding positive ions are liberated into the air and may be captured by falling raindrops. The greater the point-discharge current the greater the number of ions and so the more positive charge can be acquired by the raindrops. The greater the rate of rainfall, the more, and the larger are the raindrops so that again we get a greater rain current, but this has a limit when the rain can capture all the point-discharge ions: Simpson (1949) gave this general theory to explain his results, and later more details were worked out.

7. Acquisition of charge by falling drops

Wilson (1929) proposed a theory according to which water drops falling in a field could, in suitable circumstances, acquire a charge of one sign from the ions in the air. This theory was originally applied to conditions within a cloud and Wilson used it to suggest a mechanism for the building up of the charges in a thunder cloud. But, as Wilson himself also realized, the same theory can be applied to drops falling below a cloud. Wilson's theory was worked out in detail by Whipple and Chalmers (1944) and applied by Chalmers (1951) to the results of Simpson (1949) and of Hutchinson and Chalmers (1951).

Wilson's theory can be explained simply as follows: a drop in a field is polarized, since it is a conductor. If the potential gradient is negative, the top of the drop has a positive charge and the bottom a negative. If the drop was at rest, it would receive charges of both signs from the ions in the air, negative at the top and positive at the bottom; but when the drop is falling it may be falling more rapidly than the motion downwards of the negative ions in the field, so that it catches up with the negative ions, which are then repelled by the negative charge in the lower part of the drop and never reach the upper part of the drop unless the drop has a resultant positive charge. The positive ions still reach the bottom of the drop and so the drop acquires a positive charge in a negative potential gradient. When there are positive ions alone present rather than ions of both signs, the charge acquired by the drop is greater by a factor of over 5.

When point discharge is occurring, as is the case in those measurements where this theory has been applied, then there will be a large excess of ions of one sign. If the rain obtains its charge from the point-discharge ions, then we would expect the connection between rain current and point-discharge current as found by Simpson, and this ratio would increase with the rate of rainfall, at first proportionately, but approaching a limit for high rates of rainfall when the whole of the point-discharge current would be brought back by the rain. This qualitative picture by Simpson (1949) was made more precise by Chalmers (1951) and gives a reasonably good account of the charging of average drops during steady conditions. However, the picture does require that the potential gradient increases quite rapidly with height, a result for which observational evidence is still lacking. Smith (1955) was able to extend similar calculations to the case of drops of different sizes falling under similar conditions, but also was unable to explain divergences between similar drops. However, as a first approximation, it does appear that the capture of ions of the point-discharge current gives an account of the relation between precipitation current

and point-discharge current, and hence of the charge on precipitation from cumulo-nimbus clouds.

8. The mirror-image effect

Simpson (1949) found that, in stormy weather when conditions are not steady, the curves relating precipitation current with time and point-discharge current with time appear to be mirror images of one another, so that the relation between the two currents is not confined to steady conditions. Sivaramakrishnan (1957) found that the change of sign of the precipitation current synchronizes more closely with the change of sign of the potential gradient at the ground than with that of the point-discharge current.

A raindrop takes several minutes to fall from the cloud to the ground, so it would seem that the mirror-image effect, with simultaneous changes of both currents, would indicate that the raindrop acquires its charge fairly close to the ground. But if, as is shown by the calculations of Chalmers (1951) and Smith (1955), the drop acquires its charge by capture of ions according to Wilson's process, the potential gradient close to the ground is too small and the drop must have acquired its charge at a higher level where the potential gradient would be greater. This discrepancy has not yet been solved.

9. Rain charges in low potential gradients

It is found that precipitation is charged even when the potential gradient is too small to give point discharge. The theory of Wilson (1929) does give charging to drops when ions of both signs are present, but calculations show that the observed charging is much greater than could be accounted for in this way; since ions of both signs are present, there is no space charge as in the case of point-discharge ions, and thus no increase of potential gradient with height.

There is, however, a relation between the precipitation current and the potential gradient, giving a general inverse relation although the current is found to be zero, not for zero potential gradient but for a potential gradient which is something like the undisturbed fine-weather potential gradient.

None of the attempts to explain these results is really satisfactory, and the question of the difference between currents in rain and snow is probably important.

10. Results for snow

Results for snow show marked differences from those for rain. In turbulent conditions, snow is most usually found to be positively charged, but snow in quiet, steady conditions is almost always negative (Chalmers 1956). Snow also shows the inverse relation between precipitation current and potential gradient, but whereas the line relating rain current with potential gradient for rain lies almost wholly above the axis and has an intercept of positive potential gradient, that for snow lies below the axis and has an intercept of negative potential gradient.

It would appear that there are two possible explanations of the difference between rain and snow; either there is a charge-separation process which takes place when the snowflakes melt to form rain, or else there are charge-separation processes close to the ground giving different effects for liquid and solid precipitation.

For rain in this country from nimbo-stratus clouds the meteorological evidence is very strong that it has originated by the process suggested by Bergeron (1933), so that every raindrop has been a snowflake at an earlier stage of its history. If, therefore, there is any charging process for snowflakes in the cloud, all raindrops will have experienced this process before melting.

11. Effects of splashing

Smith (1955) thought that the positive charge on rain and the negative potential gradient might be accounted for by the separation of charge on splashing (Lenard 1892). But the measurement of raindrop charges by methods in which the charge is measured before the drop can have splashed show that this simple picture cannot be correct, and that raindrops usually have a positive charge before they can splash.

There is, however, evidence for the production of a negative space charge in the lower air during rain; Kelvin (1860) with a tower in Glasgow and Chauveau (1900) at the Eiffel Tower found occasions, during rain, when the potential gradient remained positive at the top of the tower but changed to negative at the ground, results which can be explained only if there is a negative space charge, which could be produced either by splashing at the ground or by some process in the lowest regions of the air.

Simpson (1915) made the suggestion that gusty winds near the ground might cause impacts between drops, giving rise to positive charges on the rain and negative in the air, but no direct evidence for this process is yet available.

12. The snow cloud

The difference between quiet snow and quiet rain is very marked, since snow is usually negatively charged and rain positively. An obvious suggestion is that the difference is to be ascribed to different effects at, or close to, the ground. If gusty winds produce impacts and breaking, snowflakes might get negative charges and raindrops positive and so the results could be explained.

But consideration of the currents through a cloud, from which there is quiet snowfall, shows that the acquisition of charge close to the ground cannot explain the results. The total current reaching the ground during quiet snowfall is negative and, if we are able to consider conditions as steady, then there must be a current negative downwards at all levels. Below the cloud, if we consider that the falling snowflakes are uncharged until they approach the ground, the current must be carried by the upward movement of positive charge released near the ground where the snow gets its charge. Thus between the cloud and the ground, there must be a positive space charge, and so the potential gradient just below the cloud is more negative than that at the ground. Since the potential gradient near the ground is about as often negative as positive during quiet snow it follows that the potential at the base of the cloud is usually negative; in the cases where the potential gradient at the ground is positive, the snow has a high negative charge and hence there would be a large positive space charge, again leading to the conclusion of a negative potential at the base of the cloud.

Since the electrosphere is at a positive potential of average value about $2{\cdot}9 \times 10^5$ volts, a current bringing negative charge down above the cloud could occur only if the top of the cloud has a potential higher than this value.

Thus, if the snow gets its charge close to the ground, we are led to the conclusion that there must be a potential

difference within the cloud of magnitude appreciably greater than the fine-weather potential difference between the electrosphere and earth. Such a potential difference would require some charge-separation process acting within the cloud and we are therefore forced to abandon the idea that the only charge-separation process is near the ground.

If we need a process of charge separation within the snow cloud, we can consider whether this may be the most important, if not the only, charge-separation process at work. The snow would then acquire its negative charge in the cloud, leaving behind a positive charge which is dissipated to the electrosphere. It would seem possible to explain the general relation between precipitation charge and potential gradient from the idea that the larger the charge on the snow, the larger the charge left behind and so the larger the potential gradient which this charge produces at the ground.

13. The rain cloud

As already mentioned, it is fairly certain that the rain which falls from nimbo-stratus clouds in this country has been formed by the Bergeron mechanism and so has been in the solid form in a part of the cloud. It is difficult to avoid the conclusion that the charge-separation process in the cloud, which, as we have seen, is responsible for the charge on snow, will operate also in the appropriate part of the rain cloud. But, since the rain is found to carry a positive charge, there must be a second charge-separation process operating where the precipitation is liquid, or in the process of melting.

In the rain cloud, the total current is positive downwards and, since this must be true above the cloud, it follows that the potential of the top of the cloud is considerably lower when the precipitation is rain than when it is snow; this means that the negative charge, released when the rain acquires its positive charge, must be carried up to the top of the cloud. In the upper part of the cloud, in spite of the snow charge-separation process which brings negative charge downwards, the resultant current must be taking negative charge upwards, if there is a steady state.

14. The charge-separation processes

The above discussion has shown that there are two charge-separation processes in the continuous rain cloud: an upper one, at temperatures below the freezing-point, in which negative charge is moved downwards, and a lower one, in which positive charge moves downwards.

This may be compared with the results for thunder clouds where it is well known that there are two regions of charge separation. The directions of charge movement and the temperatures are the same for the continuous rain cloud as for the thunder cloud, but whereas in the thunder cloud the upper process is more important, in the quiet rain cloud the lower process predominates.

It may well be suggested that the same processes operate, though to different degrees, in both types of cloud. There are various theories which have been put forward to account for charge separation in thunder clouds (see Chalmers 1961) and the only contribution that the present discussion can make to the problem is that if it is accepted that the two types of cloud contain the same upper charge-separation process then the association of this process with the presence of precipitation in the solid form is more clearly shown.

Smith (1951) suggested that, in a few cases where Simpson (1949) found little charge in quite heavy rain, it might be that the rain originated from the coalescence mechanism and had never been in the solid form. If this were so, it would be an argument in favour of the lower process being connected with melting rather than some other process with rain. Dinger and Gunn (1956) have found a separation of charge of the correct sign to occur when ice, with air enclosed, melts and it may well be that it is this process which gives the lower positive charge in both types of cloud.

15. Mirror-image effect in quiet precipitation

In quiet precipitation the current and the potential gradient often appear as mirror images, although Ramsay and Chalmers (1960) found that the positive maximum of precipitation current did not quite coincide in time with the negative maximum of potential gradient, but differences of a few minutes occurred, with in some cases one maximum being first and in other cases the other. Also it was found that the 'mirror' was not the zero value.

The explanation given earlier of the relation between precipitation current and potential gradient imply that the charge separation originates at cloud level. But a recognizable mirror-image effect would suggest charge separation at a much lower level, and this discrepancy has not yet been resolved.

16. Conclusion

Contrary to what might have been predicted, it appears that precipitation does not play an important part in the electrical balance of a storm cloud; in quiet precipitation its part must be greater. Considering the whole world, the present evidence is that precipitation assists the fine-weather current, rather than tending to maintain it.

Presumably because of secondary processes, the precipitation current below a storm can give little indication of the charge-separation processes within the cloud.

While the general charging of precipitation can be explained reasonably satisfactorily, there are many features concerned with the mirror-image effect and with the variations between individual drops which are yet unexplained.

It is clear that more observations are required, under different conditions, before it can be hoped that a complete understanding of the electricity of precipitation will be reached.

References

ADAMSON, J., 1960, *Quart. J.R. Met. Soc.*, **86**, 252.

ARABADJI, V. I., 1959, *Akad. Nauk. SSSR*, **127**, 298.

BERGERON, T., 1933, *Gen. Ass. Int. Un. Geod. (Lisbon), (Proc. 5th Assembly UGGI)*, p. 156.

CHALMERS, J. A., 1951, *Quart. J. R. Met. Soc.*, **77**, 249.

—— 1956, *J. Atmos. Terr. Phys.*, **9**, 311.

—— 1961, *Bull. Inst. Phys. and Phys. Soc.*, in the press.

CHALMERS, J. A., and LITTLE, E. W. R., 1947, *Terr. Mag. Atmos. Elect.*, **45**, 451.

CHALMERS, J. A., and PASQUILL, F., 1938, *Proc. Phys. Soc.*, **50**, 1.

CHAUVEAU, B., 1900, *Ann. Bur. Mét. Fr.*, **5**, 1.

DINGER, J. E., and GUNN, R., 1956, *Terr. Mag. Atmos. Elect.*, **51**, 477.

ELSTER, J., and GEITEL, H., 1888, *Met. Z.*, **5**, 95.

GISH, O. H., and WAIT, G. R., 1950, *J. Geophys. Res.*, **55**, 473.

GSCHWEND, P. P., 1922, *Jb. Radioakt.*, **17**, 62.

GUNN, R., 1949, *Rev. Sci. Instrum.*, **20,** 291.
HERATH, F., 1914, *Phys. Z.*, **15,** 155.
HUTCHINSON, W. C. A., and CHALMERS, J. A., 1951, *Quart. J. R. Met. Soc.*, **77,** 85.
ISRAËL, H., 1953, *Geofis. Pur. Appl.*, **24,** 3.
KELVIN, LORD, 1860, *Papers on Electricity and Magnetism* (London: Macmillan), p. 208.
LENARD, P., 1892, *Ann. Phys., Lpz.*, **46,** 584.
MCCLELLAND, J. A., and NOLAN, J. J., 1912, *Proc. R. Irish Acad. A*, **29,** 81.
RAMSAY, M. W., and CHALMERS, J. A., 1960, *Quart. J. R. Met. Soc.*, **86,** 530.
SCHONLAND, B. F. J., 1928, *Proc. Roy. Soc.* A, **118,** 252.
SIMPSON, G. C., 1915, *Phil. Mag.*, **30,** 1.
—— 1949, *Geophys. Mem., Lond.*, **84,** 1.
SIVARAMAKRISHNAN, M. V., 1957, *Indian J. Met. Geophys.*, **8,** 379.
SMITH, L. G., 1951, *Quart. J. R. Met. Soc.*, **77,** 683.
—— 1955, *Quart. J. R. Met. Soc.*, **81,** 23.
STERGIS, C. G., REIN, G. C., and KANGAS, T., 1957, *J. Atmos. Terr. Phys.*, **11,** 83.
WAIT, G. R., 1950, *Arch. Met., Wien, A*, **3,** 70.
WHIPPLE, F. J. W., and CHALMERS, J. A., 1944, *Quart. J. R. Met. Soc.*, **70,** 103.
WILSON, C. T. R., 1929, *J. Franklin Inst.*, **208,** 1.
WORMELL, T. W., 1930, *Proc. Roy. Soc. A*, **127,** 567.

An induction furnace to attain temperatures above 3000°C in controlled atmospheres

by L. C. F. BLACKMAN,* Ph.D., P. H. DUNDAS, Ph.D., A. W. MOORE, B.Sc., and A. R. UBBELOHDE, M.A., D.Sc.(Oxon), F.R.S., Department of Chemical Engineering and Chemical Technology, Imperial College of Science and Technology, London, S.W.7

MS. received 29th December 1960

Abstract

An induction furnace operating at 500 kc/s is described which is capable of providing temperatures in excess of 3000° C at high vacuum or in any suitable gas mixture at pressures up to 8 atmospheres. The work coil is mounted inside the furnace chamber, which can then be designed as a pressure vessel and constructed of electrically conducting steel. By this technique close coupling is provided between the work coil and the susceptor crucible: if desired, a heat treatment cycle to 3000° C can be completed in less than one minute. To avoid any contamination, the furnace is thermally uninsulated, so that all the heat dissipated in the crucible is lost by radiation, limiting its total surface area to about 20 cm² for temperatures of 3500° C with a 25 kW generator.

Application of the theory of induction heating shows that the effective resistance of the susceptor crucible can be increased by several methods, thus improving the heating efficiency of the furnace. The theory may also be applied for determining the resistivities of various susceptor materials at very high temperatures.

1. Introduction

The growing importance of graphite as a high temperature ceramic and as a solid with special electronic properties has led to the demand in this laboratory for a high temperature furnace in which the specimens can be kept free from any chemical contamination and can be heated to over 3000° C. To meet this need a furnace has been designed which can be controlled over the temperature range 1500–3100° C and which can be operated either under high vacuum conditions or up to a positive pressure of 8 atmospheres in any suitable gas mixture. The furnace is both clean and simple to use; compared with a conventional carbon resistance furnace it gives a saving in time of about 5 hours for any one heat treatment cycle. Also, it can be used for rapid heating and chilling at rates above 200 deg C/s.

The design problems encountered are discussed in detail below. Preliminary considerations of the materials available and the method of heating reveal certain relevant advantages and disadvantages. The materials give little choice since the only solid which could at present compete against graphite above 3000° C is tungsten, a metal which is difficult to fabricate and is not suitable for a crucible material for heat treatment processes involving graphite. The superiority of graphite prepared by pyrolysis of gaseous hydrocarbons over polycrystalline graphite prepared by coking is offset by the greater ease with which the polycrystalline material is at present obtainable. Detailed criteria in the choice between these two graphites are discussed later in this paper.

The type of furnace and the method of heating also requires some consideration. The method involving the simplest equipment—that of passing a large current through a carbon resistance tube—suffers from the disadvantages: (1) a period of about 5 hours is required to complete a single short heat treatment at 3000° C; (2) samples to be heated are liable to contamination by volatile impurities in the resistance tube; (3) the furnace cannot conveniently be operated at pressures greater than one atmosphere—as is desirable in order to reduce evaporation of carbon at the highest temperatures; and (4) above about 2600° C (Hove 1958) the breaking stress of graphite falls off markedly so that around 3000° C stresses placed on the heating element by the fixed electrodes tend to fracture the tube. By using induction heating, a furnace may be constructed with the advantages: (i) large amounts of energy can be concentrated in small volumes with the result that a heating cycle to 3000° C can be completed, if necessary, in less than one minute; (ii) contamination of the furnace atmosphere can be reduced to a negligible amount; (iii) heating can be conveniently carried out at pressures appreciably higher than one atmosphere; and (iv) since the graphite heating element and susceptor is not attached to any electrodes mechanical failure can be avoided.

2. Induction furnaces

Induction furnaces operating above 3000° C have been previously described. Zaer in 1935 used such a furnace for measurement of the sublimation point of graphite (3500 ± 10° C at one atmosphere pressure). In an improvement of this design, a laboratory furnace was constructed in which a graphite crucible with a volume of 150 cm³ could be heated to 3000° C with only a 5 kw generator (Ribaud 1950). The induction furnace has also been developed in industry for use up to 3600° C (Chestnut 1953).

However, all these induction furnaces have certain serious disadvantages when stringently controlled conditions are required since they rely on carbon black for thermal insulation. This imposes the limitation that the furnace cannot be evacuated or conveniently used at pressures much greater than one atmosphere, and introduces a high possibility of chemical contamination of the specimen from the gaseous impurities evolved when the carbon black becomes heated. The loading and dismantling of these furnaces is also both dirty and tedious, and they take a considerable time to reach steady operating conditions.

The furnace described in this paper can be operated above

*Now at British Railways Research Department, The Avenue, Muswell Hill, London, N.10.

3000° C and avoids the above limitations. Moreover, it can be used for extremely rapid heat treatments on graphite and certain refractory carbides, borides and nitrides and for many other operations at high temperature. One disadvantage in this method of heating is that the induction generators have a maximum overall power efficiency of only about 40%. This, however, is acceptable to industrial application as it is nullified by the great increase in operational efficiency. At a given power input the maximum temperature depends solely on the surface area of the heated crucible since heat losses at these very high temperatures are entirely by radiation. The radiation losses can be reduced by employing radiation shields, for example, of polished molybdenum.

3. Choice of generator for heating small graphite susceptors

The choice of a high frequency generator follows from consideration of the basic theory of induction heating which is briefly reviewed. When an alternating current is induced in a cylindrical conductor, heat is produced only in a surface (skin) layer provided the frequency of the current is sufficiently high. The skin thickness δ is a function of the resistivity ρ of the load material, and the frequency f of the induced current as defined by the formula

$$\delta = \frac{1}{2\pi}\left(\frac{\rho}{f}\right)^{\frac{1}{2}}. \qquad (1)$$

All quantities are in electromagnetic c.g.s. units.

Ribaud (1950) showed that for optimum results the minimum frequency which can be used is a function of the resistivity and the diameter d of the material to be heated.

$$f_{\min} = \frac{1 \cdot 8\rho}{d^2}. \qquad (2)$$

Clearly, the maximum temperature which can be obtained in such a conductor when used as a furnace element depends on the amount of heat lost by radiation. For this reason, if radiation shields are not used it is necessary to construct a susceptor crucible with small surface area. In order to produce very high temperatures around 3500° C with a 25 kw generator operating at 80% coil efficiency the maximum total surface area of the crucible should not exceed about 20 cm^2. This corresponds to a crucible whose height and diameter are about 2 cm. From Eqn (2) the minimum frequency to be used is 450 kc/s. Several workers have employed these very high frequencies for heating small susceptors to 2600° C without any radiation shields (Smyth *et al.* 1951, Harris and Jenkins 1959).

For the present work a valve-oscillator generator operating at 500 kc/s has been chosen. It is the Redifon type IH46 Mk III capable of delivering up to 25 kw of radio-frequency power.

4. Mechanical design

The furnace assembly is shown in Fig. 1 and consists of a seamless stainless steel cylinder A of 8½ in. diameter closed at both ends by thick water-cooled brass flanges B. The system is made vacuum tight by the use of neoprene 'O' ring seatings C as shown.

This design incorporates the special feature of mounting the work coil assembly D within the vacuum chamber (Smyth *et al.* 1951) which allows a much closer coupling to be employed between the load and the coil, but introduces the problems of electrical insulation and possible electrical breakdown of the gas. The problem of ensuring electrical insulation through the wall of the furnace chamber is overcome by the use of a neoprene block shown in Fig. 2. The problem of sparking through the gas is discussed later.

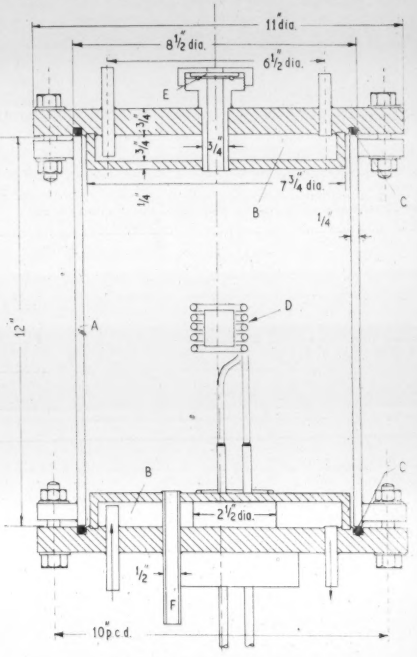


Fig. 1. Sectional view of induction furnace.

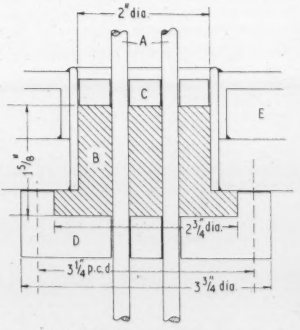


Fig. 2. Practical detail of vacuum-tight power lead insulation. A, high frequency leads; B, neoprene vacuum-tight insulation; C, pyrophyllite radiation shield; D, ebonite cover plate; E, water-cooled furnace base.

The temperature of the graphite crucible was measured optically through a silica window E mounted in the top brass flange B and using a high precision (±0·75%) potentio-

metric optical pyrometer (Evershed and Vignoles). A glass prism was placed directly above the silica window to deflect the light rays in a horizontal plane for ease of viewing. It can be shown that the correction for light absorption to be made on any measured value is given by the equation $\Delta T = 1{\cdot}3 \times 10^{-5}\, T^2$ (cf. Kingery 1959) where T is the observed temperature in degrees Celsius.

The furnace chamber which was evacuated via tube F was designed as a pressure vessel to withstand 8 atmospheres pressure with a safety factor of 4 (British Standards Code 1500: 1949). All relevant dimensions are shown in Fig. 1.

5. Operation under anaerobic conditions

High-frequency sparking has proved to be troublesome in certain cases. Practical work carried out by Holst and Koopmans (1918) and Klemperer (1938) on the sparking potential of nitrogen and the helium group of gases show that under the same voltage and pressure conditions the sparking distance of dry argon is 25 times greater than that of dry nitrogen. However, nitrogen cannot be used because it reacts with graphite at high temperatures to form cyanogen. Margrave (1955) has collected high temperature thermodynamic data and from his values it can be shown that for one atmosphere of nitrogen the reactions are as follows:

	partial pressure of cyanogen (atm.)	
	1730° C	2230° C
$2\,C_g + N_{2g} \leftrightharpoons (CN)_{2g}$	10^{-5}	$3{\cdot}34$
$2\,C_s + N_{2g} \leftrightharpoons (CN)_{2g}$	$\sim 10^{-12}$	8×10^{-4}

where g and s refer to gaseous and solid carbon respectively. Thus above about 2000° C in nitrogen, cyanogen is readily formed by gaseous reaction so that the graphite crucible then has a very short life. This imposes the use of argon as the most suitable choice of inert gas.

The minimum convenient distance between any high voltage point and any earthed point in this present design was about 1 cm. This means that when the furnace is used with argon at one atmosphere pressure the approximate maximum voltage difference that can be established without sparking is 1500 v (according to the data for d.c. peak voltage given in Fig. 3). However, in practice, using high frequency, appreciable sparking is observed above 1100 v. This was acceptable for coils with low inductance of diameter not exceeding 3·2 cm and effective working has been obtained with these coils up to 3100° C.

When the size of the crucible is increased, a larger coil has to be used with the result that in one atmosphere of argon maximum loading cannot be achieved before the peak voltage across the coil rises to a level where sparking occurs. Without using radiation shields, the maximum operating temperature was found to be about 2200° C.

At high vacuum (about 10^{-3} mm Hg) the furnace can be used up to 2200° C at which temperature the graphite begins to evaporate appreciably (Kelly 1935). At intermediate pressures around 1 mm the present design of the furnace cannot be used because of sparking breakdown (cf. Fig. 4). However, at pressures above one atmosphere the sparking distance, which decreases inversely with increasing pressure, is again sufficiently low to permit useful operation.

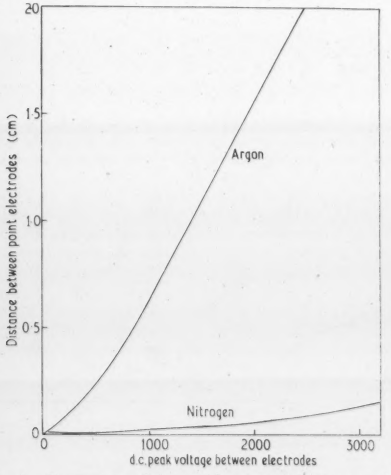


Fig. 3. Comparison of the sparking potential between point electrodes in argon and in nitrogen.

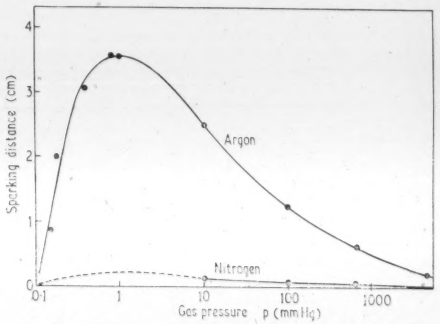


Fig. 4. Effect of gas pressure on the sparking distance between point electrodes in argon and in nitrogen.

6. Design of work coil for induction heating

6.1. *Nomenclature*

All units are electromagnetic c.g.s. units except where stated.

A, surface area of crucible; d, diameter of crucible; D, inside diameter of work coil; e, thickness of crucible wall; e_{opt}, optimum thickness of crucible wall; E, peak voltage across work coil terminals; f, frequency of induced current; h, height of crucible; H, height of work coil; i, current induced in load; I, root mean square current in work coil; I_A, maximum current available from generator at any given work coil inductance; I_R, work coil current required to dissipate maximum power in the load and the work coil winding; k, ratio e/e_{opt}; k_e, field factor (ratio of square of magnetic field intensity obtained inside a short solenoidal work coil to that obtained in an infinitely long work coil of the same diameter) ($k_e \not> 1$); k_f, geometrical shape factor (ratio of power dissipated in a short cylindrical load to that calculated by treating the load as having infinite length) ($k_f \not< 1$); K_1, coil shortness factor; L, inductance of work coil and associated leads (μH); n, number of turns in the work coil winding per cm; r, resistance of load; T, temperature of crucible (° C); T_a, absolute temperature of crucible (° K); W, energy lost by

radiation; W_c, energy dissipated in a long cylindrical load inside a coil of large H/D; W_e, energy dissipated in the work coil; W_s, energy dissipated in a short solid cylindrical load placed inside a short coil; δ, thickness of skin layer through which induced eddy current flows; ε, heat emissivity of cylindrical load; σ, Stefan–Boltzmann constant ($5{\cdot}67 \times 10^{-15}$ kw/cm^2 deg K^4); ρ, resistivity of cylindrical load or crucible; ρ', resistivity of work coil winding.

6.2. *Theoretical considerations*

In the thermally uninsulated induction furnace, because of the high temperatures used, the heat lost from a cylindrical crucible placed at the centre of a solenoidal work coil may be assumed to be totally due to radiation, and since the temperature of the crucible is so high with respect to the surroundings (when radiation shields are not used) the temperatures and emissivities of these bodies may be neglected in determining the heat loss. Thus

$$W = \sigma\varepsilon A T_a^4. \tag{3}$$

The emissivity of polycrystalline graphite at the temperatures used can be taken as 0·78 (Campbell 1956).

The maximum output of the generator is 25 kw and the best heating efficiency will be achieved when the current required to dissipate this energy in the load and in the winding of the work coil just matches that which is available at the inductance of the work coil used. As will be seen later, such a condition is impossible to achieve using small crucibles of the dimensions described above, but with larger crucibles (described later) a coil can be designed to give optimum loading. In general it is desirable to make the dimensions of the work coil as small as possible in order to reduce its inductance and thereby increase the current available for heating the load.

Figure 5 shows the schematic diagram of the furnace work

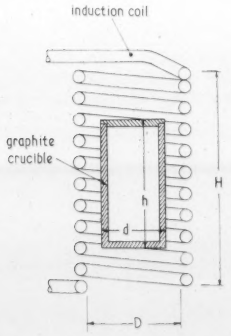


Fig. 5. Basic arrangement for induction heating.

coil and its cylindrical load, and the dimensions referred to in subsequent design calculations.

The following design work is based on theoretical and practical relationships derived by Ribaud (1950).

The magnetic field intensity induced inside a long solenoidal work coil C is $4\pi nI$. For an alternating current where the flux varies sinusoidally the effective field is the root mean square of the maximum field.

In the case of a long solenoid and long axial load the expression for the energy dissipated in the load may then be derived as follows:

$$r = \rho \frac{\text{effective length}}{\text{effective area}} = \frac{\pi d\rho}{h\delta} \quad \text{and} \quad i = n_1 hI$$

thus

$$W_c = \frac{\pi d\rho h}{\delta} n_1^2 I^2. \tag{4}$$

On eliminating δ,

$$W_c = 2\pi^2 n_1^2 I^2 hd(\rho f)^{\frac{1}{2}}. \tag{4a}$$

Similarly, the power dissipated in the copper winding is

$$W_e = 2\pi^2 n_1^2 I^2 HD(\rho' f)^{\frac{1}{2}}. \tag{5}$$

For a hollow cylinder of wall thickness $e > \delta$ the expression for W_c remains the same, but as the wall thickness decreases to values less than δ, the energy absorbed by the load increases toward a maximum at an optimum thickness given by $e_{opt} = 2\delta^2/d$ before decreasing to zero as e decreases to zero. It can be shown that the energy absorbed by such a thin cylinder is given by

$$W = \frac{d}{4\delta}\left(\frac{1}{2k} + \frac{k}{2}\right)^{-1} W_c. \tag{6}$$

The above expressions apply rigorously to the heating of long cylindrical loads coaxial with work coils of large H/D. However, in practice, one often uses loads and work coils of such dimensions that neither h/d nor H/D are large. In the case of a solid cylindrical load of small h/d current is induced not only in a thin layer of skin at the periphery, but also at both of the ends of the cylinder to a depth of δ, since the lines of magnetic field intensity converge toward the axis of the load at the ends. If it is assumed that the height of the zone in which the field lines tend towards the centre is proportional to the crucible diameter d, then the energy dissipated in the ends is proportional to d^2, and it can be shown that the 'geometrical shape factor' would be equal to $1 + kd/h$. Experimental measurements on many samples (Ribaud 1950) show that the value of k in this expression is 1·15 and

$$k_f = 1 + 1{\cdot}15d/h. \tag{7}$$

This means that for a cylinder with $h/d = 1$, there is 2·15 times as much energy dissipated in the cylinder as given by Eqn (4).

The above discussion still assumes that the work coil is of infinite length (i.e. large H/D). However, such an assumption is not normally permissible; for a 'short' coil the field produced at the centre is less than $4\pi nI$ because the lines of force are not parallel to the coil axis. Thus it is necessary to multiply the expression for energy dissipated in the coil (Eqn (4)) not only by the 'geometrical shape factor' but also by a 'field' factor. This factor is the ratio of the square of the magnetic field intensity obtained inside a short solenoidal work coil to that obtained in an infinitely long work coil of the same diameter. The field factor at all points inside a short solenoid of any given H/D can be calculated and integrated throughout the volume to give the appropriate value for any cylindrical load axially placed. Values presented by Ribaud (1950) for the particular case of a solenoid of $H/D = 1$ were used in the following theoretical design considerations, since the work coils used in the present work

had this same dimensional ratio. The energy received by a solid crucible inside a coil of finite dimensions is thus given by

$$W_s = W_c k_f k_c. \tag{8}$$

Equating the expression for W_s and W_e to 25 kw, the maximum output of the generator

$$W_c k_f k_c + W_e = 2{\cdot}5 \times 10^{11}. \tag{9}$$

The expression for the inductance of the loaded work coil is (Vaughan and Williamson 1945)

$$L = 0{\cdot}01 n_1^2 K_1 (HD^2 - hd^2) + 0{\cdot}01\frac{d^2}{h} + \text{leads inductance.} \tag{10}$$

The coil shortness factor K_1 is a function of Nagaoka's constant (Vaughan and Williamson 1945) and the diameter of the coil and its load. By approximating for Nagaoka's constant in terms of a linear function of D/H, the expression for K_1 is obtained:

$$K_1 = 1 - \frac{0{\cdot}35}{H}\left(D - \frac{d^2}{D}\right). \tag{11}$$

7. Optimum dimensions of work coil for crucible of given dimensions

To attain very high temperatures in the induction furnace, crucibles have been used of diameter 2·0 cm and height varying from 2·4 to 2·7 cm. The following typical examples illustrate the results to be expected from theory (applying Eqns (1), (4), (5), (7), (8) and (9)) when coils and crucibles of various dimensions are used in the present design of induction furnace.

7.1. *Crucible 2·0 cm diameter × 2·4 cm high (wall thickness $e > \delta$)*

If the resistivity of the graphite forming this crucible is 10^6 c.g.s. the current I_R required to dissipate the available 25 kw from the generator is at least 500 amps. Fig. 6 shows the maximum available current from the generator as a function of the loaded work coil inductance. To obtain a work coil current of 500 amps would require a work coil inductance of only 0·3 μH. Such a condition is not feasible because in the present design the work coil leads inductance is always about 0·4 μH. However, it has been found unnecessary with these small crucibles to dissipate the maximum power to obtain temperatures in excess of 3000° C. Accordingly a work coil of 3·1 cm height (4 turns of ¼ in. copper tubing) and 3·2 cm diameter was constructed. Such coils are made by turning the required length of copper tubing around a cylindrical rod of suitable diameter. The copper is then annealed by heating to a red hot temperature and quenching in water. This procedure permits one to fit the new coil more easily into position in the furnace.

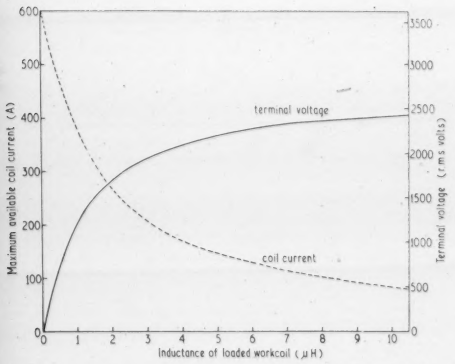


Fig. 6. Output characteristics of induction generator.

The loaded work coil inductance calculated from Eqns (10) and (11) is 0·70 μH, and from Fig. 6 the maximum available coil current is 400 amps. At this current the maximum power which can be dissipated in the load (from Eqn (9)) is 14·5 kw ($\rho = 10^6$). If the graphite has a resistivity of $1{\cdot}2 \times 10^6$ c.g.s. the maximum power is 15·5 kw.

7.2. *Crucible 2·0 cm diameter × 2·7 cm high (wall thickness $e < \delta$)*

For such a crucible with wall thickness $e = 0{\cdot}15$ cm and $\delta = 0{\cdot}23$ cm ($\rho = 10^6$) and the same work coil described in § 7.1, the maximum power dissipated in the load is 17·0 kw.

7.3. *Crucible 2·7 cm diameter × 3·9 cm high (wall thickness $e < \delta$)*

To permit investigation of the effect of radiation shields on the effective power input to the crucible (i.e. the temperature obtained at a given power input) this larger crucible was constructed. To accommodate the molybdenum radiation shield coaxially between the coil and the crucible, a coil of diameter 5·5 cm and height 6·0 cm was constructed. By making the wall thickness of the crucible 0·15 cm, the maximum power which can be dissipated in it is 15 kw at $\rho = 10^6$.

7.4. *Crucible of same dimensions as in § 7.3 with optimum coil*

It should be noted here that with the larger crucibles it is possible to design optimum coils that will allow 25 kw of power to be dissipated in the load and the winding. For example, in the case of the crucible just mentioned, a coil of diameter 4·5 cm and height 4·5 cm would be satisfactory for this purpose. However, its diameter would not be large enough to accommodate the radiation shield.

These preliminary calculations show that because the leads inductance is such a significant fraction of the total work coil inductance, it is not possible to dissipate the maximum available power of 25 kw in systems containing small loads. It can be seen, however, that the best heating efficiency can be obtained both when the relative space occupied inside the coil by the crucible is the largest possible and when the effective resistance of the crucible is made as large as possible (e.g. by decreasing the wall thickness to values less than δ).

8. Relative merits of polycrystalline and pyrolytic graphite as a susceptor material

In most of the experimental work carried out in this furnace polycrystalline graphite (Morgan EY9A) crucibles have been used because they are readily made from solid cylindrical rods. The method of affixing these crucibles to the furnace is illustrated in Fig. 7(*a*). The crucible support C is electrically insulated from the base of the furnace by a pyrophyllite

disk P. A graphite cylindrical rod C is supported on the pyrophyllite. The small graphite crucible is then supported on a $\frac{1}{8}$ in. rod S of polycrystalline graphite (of small diameter so that the induced eddy current near the coil is small) which

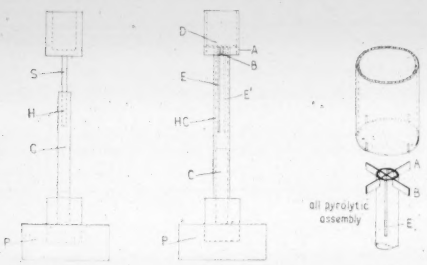


Fig. 7. Method of mounting crucibles.

fits into the base of the crucible. The $\frac{1}{8}$ in. rod fits tightly into a concentric hole H drilled into the rod C, forming a rigid support for the crucible.

The disadvantage of using polycrystalline graphite as a susceptor material is that it is generally very porous and has to be heated under vacuum to high temperatures to eliminate occluded gases. Even after heat treatment for 30–40 minutes under vacuum at 2200° C, when the graphite is heated in an atmosphere of argon to temperatures much above 2800° C, considerable volatile matter is evolved. In addition to contaminating the furnace, the volatile material makes the furnace atmosphere more conducive to high voltage sparking. Because of these disadvantages, the use of pyrolytic graphite as a susceptor material has been considered and a few experimental runs have been made with this material.

The only disadvantage of using pyrolytic graphite as a susceptor and support is the present relative difficulty of manufacturing the various shapes needed for this construction. Figs 7(*b*) and 7(*c*) show the technique employed in this laboratory for mounting an arrangement using all-pyrolytic graphite with the exception of the base supporting rod C which is assessed to reach a temperature not greater than 1000° C. The walls of the pyrolytic graphite crucible were formed by cracking methane at 12 cm Hg pressure and 1900° C on to a polycrystalline former of 2·0 cm diameter to a thickness of 1 mm. The deposition temperature of 1900° C was chosen since it has been shown (Blackman and Ubbelohde, unpublished) that the resulting pyrolytic graphite does not graphitize completely on subsequent heating to about 3000° C. Crystal disorder corresponding to a Franklin–Bacon (Bacon 1951) *p*-factor of about 0·1 is retained and other evidence suggests that this imparts good mechanical strength to the material.

A crucible of 2·2 cm outside diameter and 2·5 cm height was produced by this method. To form a strong support with low heat conductance for the crucible a hollow cylindrical rod HC of pyrolytic graphite was prepared which was about 1·4 cm outside diameter and 1 mm thick and 10 cm long. This was rigidly held in position by the base supporting rod C. The graphite crucible was held to this cylinder by two mutually perpendicular slotted strips A, B which fit into four slots cut in the top of the supporting cylinder and in the bottom of the crucible wall. These strips and the cylindrical bottom of the crucible D were cut from a 1 mm thick flat section of pyrolytic graphite. Two deep slots E and E′ were cut in the supporting cylinder to eliminate induced eddy currents in this section.

Preliminary results with the pyrolytic graphite show that it is entirely satisfactory as a susceptor material, its resistivity (8×10^5 c.g.s.) comparing favourably to that of the polycrystalline graphite. Although its resistivity is somewhat lower, the wall thickness of the pyrolytic graphite crucible can be made quite low, approaching the optimum, thus permitting a considerable gain in heating efficiency. It may therefore be anticipated that pyrolytic graphite will eventually replace the conventional polycrystalline graphite as a susceptor crucible, whenever the desirability of eliminating all sources of contamination in the furnace is sufficiently important.

9. Experimental results and discussion

9.1. *Heating of small crucible to above* 3000° C

Using a number of small polycrystalline graphite crucibles (2 cm diameter × 2·4–2·9 cm high) temperatures in excess of 3000° C have been achieved in this induction furnace. In general the crucibles were outgassed under vacuum for about 30 minutes at 2000–2200° C. At higher temperatures, to prevent appreciable evaporation of the carbon, one atmosphere pressure of argon was used. The voltage across the coil was determined by rectifying with a Brimar 5R4GY full wave rectifier and stored on a 0·05 μF capacitor of 5000 v d.c. rating. The voltage on the capacitor was measured on an Avometer as d.c. peak voltage across the coil.

9.2. *Effect of heat treatment on resistivity of the graphite*

The current in the work coil can be determined when the impedance of the coil is known. It can be shown that the resistive component of this impedance is negligible in proportion to the inductive reactance, so that

$$I = \frac{E}{2\sqrt{2\pi f L}}. \tag{12}$$

Data obtained with a small polycrystalline graphite crucible (2·0 cm diameter × 2·7 cm high) are presented graphically in Fig. 8. From Eqns (3), (4), (8) and (12) it is seen that the absolute temperature of a crucible should be proportional to the square root of the voltage across the coil. Theoretical curves are plotted in this figure for resistivities of $1 \cdot 0$ and $1 \cdot 4 \times 10^{-3}$ ohm cm for graphite. As is illustrated in Fig. 8, during the first heating of this crucible the experimentally indicated resistivity exceeded $1 \cdot 4 \times 10^{-3}$ until a temperature of 2800° C was reached. The results suggest that the high resistivity is due to the low degree of graphitization of this type of graphite. The apparent resistivity of the crucible decreased markedly during the 30-minute period at which the sample was held at 2800° C. When the same crucible after cooling to room temperature was reheated to 3100° C the resistivity measured was of the order of $1 \cdot 2 \times 10^{-3}$ ohm cm. This value is in good general agreement with other determinations for this type of graphite (Hove 1958), suggesting that the theory developed by Ribaud may be applied successfully to the interesting problem of determining approximate resistivities of various susceptor materials at these very high temperatures. In an experiment when the crucible was held at 3110° C a considerable amount of carbon dust was emitted from the sample, indicating the undesirability of using this polycrystalline graphite at such high temperatures.

The accuracy of the resistivity measurements by this technique depends on the accuracy with which the temperature

and heat emissivity of the load, as well as the coil current, can be determined. The emissivity has been determined quite accurately elsewhere in the case of polycrystalline graphite (Campbell 1956). Since the resistivity varies as the 8th

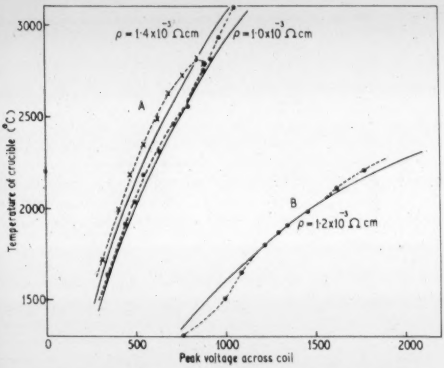


Fig. 8. Temperature of crucibles as a function of coil voltage. Correlation between theory and experimental results: A, crucible 2·0 cm dia. × 2·7 cm high, coil 3·2 cm dia. × 3·1 cm high; B, crucible 2·7 cm dia. × 3·9 cm high, coil 5·5 cm dia. × 6·0 cm high.

Theory shown by full lines, and experimental results by broken lines.

power of the absolute temperature of the load, an error of ±0·75% in optical pyrometry reading will mean an error in the estimated resistivity of about ±6%.

Our results show that the technique for determining the coil current by measuring the peak voltage across the coil and calculating the inductance of the work coil is quite satisfactory. For example, the experimentally measured peak voltage across an unloaded work coil of inductance 0·75 μH was 1470 v, compared with 1450 v calculated from Fig. 6 which presents data supplied by the manufacturers of the generator. Similarly for a larger coil (5·5 cm diameter × 6·0 cm high) the experimental peak voltage was 2520 v compared with 2560 v calculated from Fig. 6 at a calculated inductance of 2·50 H. Assuming an error of ±1% in the measurement of the coil current, an error of ±4% in the resistivity will result, since the calculated resistivity varies as the 4th power of the coil current. Any particular measurement of resistivity will thus be subject to an error of about ±10%.

9.3. *Heating of neutron bombarded pyrolytic graphite*

This furnace was used conveniently and successfully in annealing samples of pyrolytic graphite which had been subjected to damage by neutron bombardment, detailed results of which are discussed elsewhere (Blackman, Saunders and Ubbelohde, unpublished observations). Temperatures up to 2800° C in an argon atmosphere were applied to samples held in a polycrystalline graphite crucible as the susceptor material for a period of 30 minutes. The operation was completed quickly and cleanly; a negligible amount of soot particles was discharged by the crucible at this temperature.

9.4. *Heating of crucibles of larger surface area using radiation shields*

Some experimental work has been carried out using a crucible of diameter 2·7 cm and height 3·9 cm with a coil 5·5 cm diameter and 6·0 cm high. Results on a single heating of such a crucible are presented in Fig. 8 for comparison with those obtained with the smaller crucible. The experimentally measured resistivity of this crucible over the temperature range 1300–2200° C was about $1{\cdot}2 \times 10^{-3}$ ohm cm.

Preliminary work with this large crucible and coil has shown that molybdenum radiation shields increased the crucible temperature by 200–250 deg C in the temperature range 1400–1700° C. This indicates an increase of effective wattage by a factor of about 70%.

Acknowledgment

Thanks are due to the Research Council, Central Electricity Authority for their support in this work.

References

BACON, G. E., 1951, *Acta Cryst., Camb.*, **4,** 558.
CAMPBELL, I. E., 1956, *High Temperature Technology*, (London: Chapman and Hall), p. 106.
CHESTNUT, F., 1953, *Industr. Heat.*, **20,** 1690.
HARRIS, B., and JENKINS, A. E., 1959, *J. Sci. Instrum.*, **36,** 238.
HOLST, G., and KOOPMANS, A. N., 1918, *Proc. Acad. Sci. Amst.*, **20,** 1032.
HOVE, J. E., 1958, *Trans Met. Soc., Feb.*, 7.
—— 1958, *Industrial Carbon and Graphite* (London: Society of Chemical Industry), p. 501.
KELLEY, K. K., 1935, *Bur. Min. Bull.* No. 383.
KINGERY, W. D., 1959, *Property Measurements at High Temperatures* (London: Chapman and Hall), p. 12.
KLEMPERER, H., 1938, *Z. Tech. Phys.*, **19,** 270.
MARGRAVE, J. L., 1935, *J. Chem. Educ.*, **32,** 520.
RIBAUD, G., 1950, *Les Hautes Températures et leurs Utilisation en Chimie*, Chapter 9 (Paris: Masson).
SMYTH, H. T., MEINKEN, R. H., and WISNYI, L. G., 1951, *J. Amer. Ceram. Soc.*, **34,** 161.
VAUGHAN, J. T., and WILLIAMSON, J. W., 1945, *Trans Amer. Inst. Elect. Engrs*, **64,** 587.
ZAER, R., 1935, *Etude de la Vaporisation du Carbone* (Paris: Thèse Doctorat ès Sciences).

Pumping of argon, nitrogen and hydrogen in a Bayard–Alpert gauge

by B. COBIC,* G. CARTER, B.Sc., Ph.D., Grad.Inst.P., and J. H. LECK, M.Eng., Ph.D., A.Inst.P., A.M.I.E.E., Department of Electrical Engineering, The University of Liverpool

MS. received 10*th March* 1961, *in revised form* 28*th March* 1961

Abstract

The pumping of nitrogen and hydrogen in Bayard–Alpert gauges has been investigated as a function of the electrode potentials, the gauge temperature and the gas charge upon the walls. The results are compared with earlier values obtained in argon and in which the pumping mechanism is known. It is shown that ion pumping cannot account for the high pumpirg speed of nitrogen and it is concluded that mechanisms such as sputtering of the tungsten filament and the formation of tungsten nitride or the production of metastable nitrogen are unlikely to explain the observations. The pumping of hydrogen is confirmed to be due to dissociation of H_2 at the cathode with subsequent physical adsorption of H atoms at the walls, and is found to be greatly in excess of any ion pumping which may occur.

The presence of liquid nitrogen traps during experiments was found to exert a profound effect on the gauge pumping of nitrogen.

Introduction

In the production of ultra high vacuum, much use has been made of the pumping properties of ionization gauges (Alpert 1953), particularly the Bayard–Alpert (Bayard and Alpert 1950) modification of this device, generally without full knowledge of the pumping mechanism. Alpert (1953) and his colleagues and Young (1956) investigated the pumping of nitrogen and helium and Bills and Carleton (1958) observed the sorption of nitrogen but in none of these investigations were the effects of variations of important gauge parameters comprehensively studied. The reaction of hydrogen with a hot tungsten filament in the presence of a glass surface, on the other hand, has been studied by Langmuir (1915) and by Ehrlich (1961) and his associates but again the pumping properties of the Bayard–Alpert gauge for the gas have not been reported.

In this laboratory investigations of the sorption of inert gases (Carter and Leck 1961, Cobic, Carter and Leck 1961) have been undertaken in order to establish the ion pumping mechanism in these gases and it was felt desirable to conduct experiments in more active gases also to determine the importance of different processes occurring in these gases. The work described in this communication is the result of this investigation, and the essential differences between the pumping of the inert gases and the more active gases nitrogen and hydrogen are pointed out and explanations for the variations offered.

The processes which contribute towards the kinetics of the pumping in these gases is analysed in some detail and the most efficient manner of operating the gauges as pumps is also discussed.

Experimental

The experimental techniques employed in this investigation are fully described elsewhere and require only brief reiteration (Cobic, Carter and Leck 1961). Two Bayard–Alpert gauges constructed in Pyrex glass were used, one to pump the gas, the other, operated at a low emission, to record the pumping. The static volume technique was used almost exclusively for the inert gases and nitrogen, but the dynamic technique in which pumping of the gauge balanced a continuous introduction of gas was used in the case of the rapidly pumping hydrogen. The measurement of speed with the static system was made by observing the slope of the $(\log p, t)$ pump-down curve and in the dynamic system by observing the quasi-equilibrium pressure in the gauges established by a measured leak in rate. The measurement of sorbed quantity in the static system was easily effected by observing the pressure reduction while in the dynamic system this quantity was found by computing the quantity of gas which had flowed into the system plus any quantity resulting from a drop in pressure. Electron emission current and electron collector potential variation was achieved by varying component values in the emission control unit while the gauge wall temperature was varied by controlling the power input to a small oven or by immersion in a cryostatic bath.

Results

Since the pumping of argon was found to be representative of that in all the inert gases investigated, as reported earlier (Carter and Leck 1961, Cobic, Carter and Leck 1961), with the exception of helium, it is proposed to discuss this former gas only in conjunction with nitrogen and hydrogen.

Typical static pump-downs and thermal recovery in both argon and nitrogen are shown in Fig. 1 and in Figs 2, 3 and 4 are shown the relationships between speed, quantity, voltage and temperature for these gases. The most obvious difference between the two gases is that under almost all conditions argon pumps at a speed less than one fifth that of nitrogen. For the experiments on nitrogen it was found necessary to remove the liquid nitrogen refrigerant around the vapour trap and replace this with a solid CO_2–acetone mixture in order to eliminate physical adsorption effects which obscured the true pumping action. As a result of these physical adsorption effects in the presence of a liquid nitrogen trap the speed of pumping was observed to be a strong function of the gas pressure, as shown in curve 2 of Fig. 5, the implication

* Now at Institute of Nuclear Science, Belgrade.

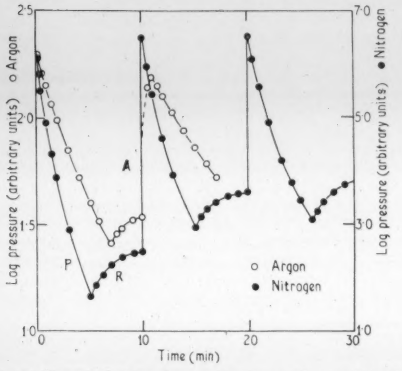


Fig. 1. Repeated pump-down and recovery characteristics in argon and nitrogen.
P, pumping region; R, recovery region; A, region in which fresh gas was admitted.

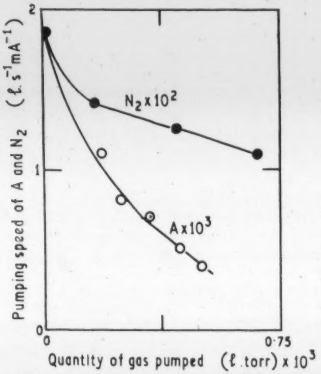


Fig. 2. The relationship between the gauge pumping speed and the quantity of gas pumped for argon and nitrogen.

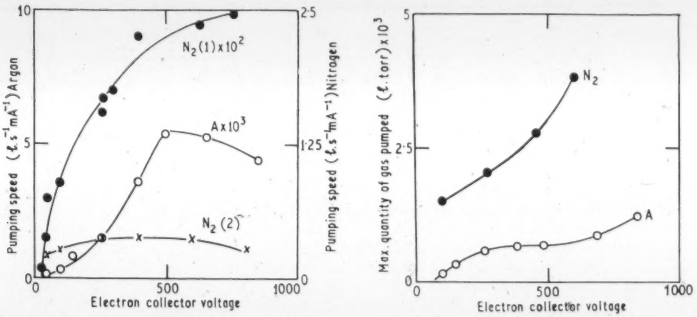


Fig. 3. (*a*) The dependence of the initial pumping speed upon the electron collector potential for argon and nitrogen. Note that curve 2 was obtained in the presence of liquid nitrogen trapping. (*b*) The dependence of the maximum quantity of gas pumped by the gauge upon the electron collector potential.

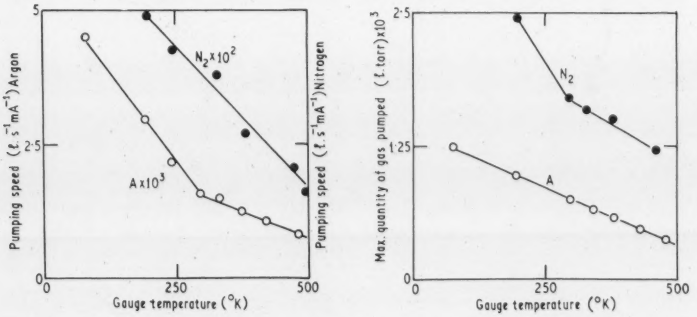


Fig. 4. (*a*) The dependence of the initial pumping speed upon the temperature of the gauge. (*b*) The dependence of the maximum quantity of gas pumped by the gauge upon the temperature of the gauge.

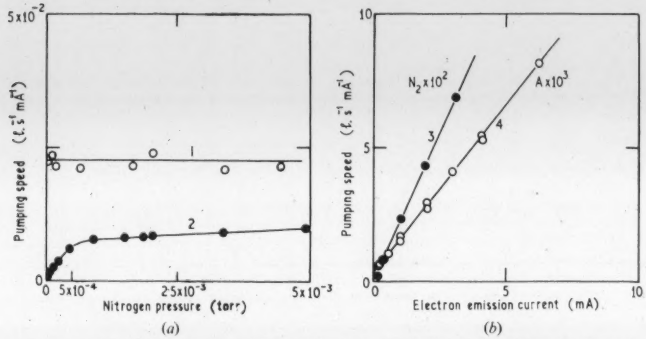


Fig. 5. (*a*) The initial pumping speed as a function of gas pressure for nitrogen. Curve 2 was obtained in the presence of liquid nitrogen trapping. (*b*) The initial pumping speed as a function of the electron emission current for argon and nitrogen. The small pumping speed at zero emission current for argon was due to a leaking vacuum valve and monitor gauge pumping.

of which is discussed more fully in the appendix, but when cooling to only −78° C was used the speed was observed to be pressure independent and linearly dependent upon electron emission current (curves 1 and 3 of Fig. 5). It is noted that the observed pumping speed of N_2 at room temperature and at an electron accelerating voltage of 250 v is of similar magnitude to that reported by Alpert (1953), Young (1956) and Bills and Carleton (1958) and the ratio of the speeds for A and N_2 are in order of magnitude agreement with results obtained with a diode type ion pump by Jenkins and Trodden (1960). The dependence of speed upon voltage in the presence of liquid nitrogen trapping is shown in Fig. 3 and the profound influence of this trap upon the pumping is further evidence. Since it was not possible to use liquid nitrogen refrigeration the thermal recovery of nitrogen could not be recorded accurately because of gross background desorption. The speed, however, fell to zero at a gauge temperature of about 450° C so that almost complete recovery could be expected by heating to 400° C. Desorption at ambient temperature was observed as in argon and as in this latter gas only 1% of the pumped gas was recovered.

In nitrogen as in the inert gases no pumping was observed if the filament was activated but no electrode potentials applied, hydrogen on the other hand was observed to pump rapidly with the filament heated only. Using the dynamic technique the pumping speed of this latter gas was determined as a function of the filament temperature and the quantity of gas pumped. The pumping action due to the heated filament was greatly in excess of any pumping due to the application of electrode potentials and the magnitude of this latter could not be determined. It is, however, probably similar to that for helium, since the two atoms have similar diffusion coefficients in glass (Norton 1954) and can be expected to penetrate the surface with similar efficiency. Thermal recovery of the pumped hydrogen could be realized by baking the glass to about 150° C indicating a range of activation energies for desorption from about 25–30 kcal/mole.

Discussion

Other investigators have advanced phenomenological theories for the sorption of nitrogen (Alpert 1953, Bayard and Alpert 1950), helium (Varnerin and Carmichael 1955) and hydrogen (Hickmott 1960) on glass, many of which make simplifying assumptions to facilitate analysis but which are not always experimentally justified. A general analysis for all gases is extremely complex and it is believed contributes little to the understanding of the phenomenon because of the many over-simplifications necessary to achieve a tractable solution. Consequently the present discussion is confined to a qualitative appraisal of the processes involved in the clean up and the way in which these contribute to the observed pumping characteristics. This analysis does involve the one basic assumption, however, that the clean up is the result of the collision of gas molecules activated by some agency, a hot metal surface, or an electron stream with discrete sites on the glass walls of the test cell. This assumption appears justified, since all experimental evidence (Young 1956, Carter and Leck 1961, Cobic, Carter and Leck 1961, Hickmott 1960) supports this belief. It is the magnitude of the various kinetic processes which occur at these sites which will determine the characteristic form of the pumping in a given gas.

In a test cell, isolated from all other sources and sinks of gas, the rate of production of activated particles will be proportional to the gas concentration and will depend upon the activating agency. Similarly the rate of bombardment of the walls will be proportional to the gas concentration and will depend upon the gas molecular speed, a function of the electrode potentials in the case of ion pumping and of surface temperatures in the case of dissociation pumping. The rate of collision of active molecules with vacant sites will be dependent upon the relative occupation of these sites. Consequently, as more sites become filled, the chance of a bombarding particle striking a vacant site diminishes, and thus the probability of clean up of a particle will decrease with increasing site occupation. On collision with a vacant site the bombarding particle may be captured or may be returned to the gas, retaining or without activity. On collision with a filled site it is possible that emission of the captured particle may occur and the bombarding particle may or may not take its place. The probability of collision of an activated particle with a filled site will increase with increasing site occupation and so the release of trapped gas

can be expected to increase as the sites fill up. Simultaneously with this release process, which is a form of gas sputtering, a second process can occur. This is the natural thermal desorption of captured particles into the gas. It has been shown that a wide variety of sites exist with a wide range of activation energies for desorption, some of which lead to a sticking time shorter or comparable to the experimental time, others which result in the almost permanent binding of the gas to the glass. Sites which have a low binding energy (less than about 15 kcal/mole) will not appear to participate in the pumping while those of high binding energy will contribute most to the sorption. Those sites whose energy leads to desorption in times comparable with the observation times will be those which lead to complexities in the pumping characteristics (i.e. sites with energies greater than about 18 kcal/mole).

Initially, when no gas is sorbed, the pumping rate will depend only upon the particle flux to the walls and the probability of sorption of a bombarding particle. In the case of rare gases it has been shown that the activated particles are ions and consequently the ratio of the clean up rate to the ion bombardment rate of the walls, will give the sticking efficiency for rare gas ions. The ion bombardment of the walls can be represented by the manometric sensitivity of the gauge and the ratio of the bombardment of the walls to that of the collector can be measured in a subsidiary experiment. From these observations it can be calculated that the ion sticking efficiency for argon is of the order 0·5 at 260 v while that for nitrogen is greater than unity at 260 v and considerably higher than this at higher and lower potentials. The implication of this will be discussed below.

As pumping proceeds sites become filled and the possibility of a bombarding particle striking a vacant site decreases, evolution from filled sites also takes place both by gas sputtering and by thermal desorption from the low energy sites. All these effects tend to diminish the pumping rate, and their relative efficiency depends upon the state of occupation of filled sites. When comparatively few sites are filled, for example by a single pump-down from a pressure of 10^{-5} torr (where the maximum quantity pumpable is of the order 0·01 l. mtorr), the first two effects tend to reduce the pumping rate only slightly. However, the third process may affect the pumping rate considerably, since even at low total site populations, the low energy sites were observed to be well populated (from desorption after pump-down measurements). Since, as the gas is pumped, the pressure falls and the rate of bombardment of the walls decreases, the rate of gas evolution from the low energy sites begins to compare with the sticking rate, consequently the observed pumping speed falls. The stage of the pump-down at which this effect becomes marked and causes a deviation from constant pumping speed will depend upon the sticking efficiency of the bombarding particles, since for a given surface coverage the desorption rate will be equalled by different bombarding rates for different sticking efficiencies. It was thus observed that nitrogen, which has an apparent sticking efficiency approximately an order of magnitude larger than that of argon, showed evidence of diminished pumping speed only at a pressure one-tenth that of argon if initial pressures were equal, while the desorption of these gases indicates similar low energy site populations. (We may note that the pressures in question here are of the order 100 × the ultimate pressure and so background gas desorption will contribute little to the speed diminution.) If at this stage the gauge is caused to stop pumping, a thermal recovery of gas having energies of activation for desorption between 20 and 25 kcal/mole is observed and if the initial desorption rate is added to the observed pumping rate just before switching off the gauge, then a true pumping speed is obtained. Experimentally, gas is then admitted to the gauge to restore the initial pressure and under these conditions the surface bombardment rate will greatly exceed the desorption rate, so that the measured pumping speed here is correct. A comparison of this speed with the calculated true speed at the end of the previous pump-down shows agreement between the two,

e.g. calculated true speed at end of first pump-down $\simeq 1{\cdot}44 \times 10^{-2}$ l. s^{-1} mA^{-1} from characteristics as in Fig. 1;
measured speed at beginning of second pump-down $= 1{\cdot}42 \times 10^{-2}$ l. s^{-1} mA^{-1};
calculated true speed at end of second pump-down $= 1{\cdot}16 \times 10^{-2}$ l. s^{-1} mA^{1};
measured speed at beginning of third pump-down $= 1{\cdot}25 \times 10^{-2}$ l. s^{-1} mA^{-1}

indicating the validity of our argument. From this discussion it is obvious that the thermal desorption processes are most pronounced at low bombardment rates. The other evolution process, the gas sputtering, which may be the result of emission by direct collision of a bombarding particle with a captured atom or by erosion of the glass (Carmichael and Trendelenberg 1958) with the attendant emission of the captured gas, on the other hand, is less effective at low bombardment rates since this process relies on collision of bombarding particles with the surface, it does increase, however, as the filled site population increases, i.e. with continued pump-down. As more gas is pumped and more sites fill, the collision frequency with empty sites decreases, and this, combined with the increased evolution, results in a drop in the pumping speed. Thus the initial speeds at each pump-down, which are essentially free from thermal desorption effects, decrease continuously. Eventually, the sorption rate will be balanced by the evolution rate when the appropriate number of sites are filled, no further net sorption will occur and the pumping speed will fall to zero. This quantity of sorbed gas is the maximum quantity which can be sorbed under either static or dynamic conditions for any bombardment rate provided the parameters ion energy and tube temperature are maintained constant. In the absence of the normal thermal desorption and evolution of background gas, it would theoretically be possible to pump all the gas from a static system provided the quantity of gas originally present was less than that required for saturation, practically the thermal desorption processes cause the effective pumping action to cease at relatively high pressures. However, experiments to be described shortly have shown how this limiting pressure can be reduced by suitable operation of the gauge. When the maximum quantity of gas pumpable has been sorbed the sticking and evolution rates are balanced, and the higher the sticking probability per bombarding particle the larger will be the total sorbed quantity. Thus the much larger quantity of nitrogen pumpable is to be expected from the apparently high sticking efficiency for this gas.

The high sticking efficiency for nitrogen in fact leads one to examine the sorption mechanism for this gas. If the sticking efficiency for argon, in which it is reasonably certain the important mechanism is the collision of ions with the walls, is computed, it is found to be of the order 0·5. Since the sensitivity of the ion gauge to nitrogen is approximately the same as that for argon (Leck 1957) and the ionization

efficiency curves (Smith 1930) for the two gases are similar, it is reasonable to believe that the ion bombardment of the walls will be similar in each case, thus leading to apparent sticking efficiencies for nitrogen of 2·5. Even if the ratio of the ion flux to the walls to that to the ion collector was underestimated in the present work (2·6 compared with between 4 and 7 quoted by Bloomer and Haine (1953) and by Young (1956)) leading to overestimates of sticking efficiencies the sticking efficiency for nitrogen is from 5 to 6 times that of argon. It is difficult to account for this increased efficiency by postulating an increased capture site density or an increased sticking probability for this gas because of its molecular size, since neon and argon, between which nitrogen is intermediate, have similar sticking efficiencies. It is thus necessary to assume that either the activated particle sorption is not entirely ionic or that the mechanism of sorption is entirely different. Considering the former possibility, Bloomer and Haine (1953) and Alpert (1953) postulated that metastable nitrogen molecules were the sorbing entities, but, while the rate of production of these can be expected to increase as a similar function of electron energy as the ionization rate (Massey 1956), it is difficult to understand how merely metastable molecules, not under the influence of electric fields, can penetrate into the energetic sites (requiring energies of activation for desorption between 25 and 50 kcal/mole).

If other mechanisms are involved the most likely will be the production of tungsten nitride on the glass surface. Langmuir (1913) has shown that at sufficiently high filament temperatures tungsten evaporates, collides with nitrogen molecules forming tungsten nitride which subsequently condenses on the glass walls. In the present work the filaments operate at lower temperatures than necessary to account for a sufficient evaporation rate, easily verified by removing electrode potentials but maintaining the filament hot, but there is likely to be a high ion bombardment of the filament. Although measurements of the sputtering of high temperature tungsten by several hundred electron volt nitrogen ions are not available, it is possible that the sputtering coefficient is of order unity. The sputtered metal may then combine with the bombarding nitrogen, and the nitride subsequently evaporate to the glass walls, or tungsten freshly sputtered to the wall may be quickly struck by a bombarding gas molecule and the nitride formed. Whichever of these may be active the pumping speed will be pressure independent as observed, whereas if collisions between sputtered tungsten and nitrogen in the gas phase were important there would be a strong pressure dependence. The magnitude of the sputtering coefficient will determine the pumping speed and it is thus possible to account for the observed speed. The speed will be expected to increase with the increasing electron collector potential as observed since the sputtering rate will increase with ion energy. If this mechanism is operative one would expect no decrease in pumping speed with increased pumped quantity since fresh tungsten would be continuously deposited, and there would be no limit to the quantity pumpable, contrary to the observations. Again, if this mechanism is important one would expect the energy of binding of the tungsten nitride to the glass to be greater than the observed maximum 50 kcal/mole, since it will undoubtedly be a chemical bond (the pumping action of nitrogen is observed to fall to zero at about 450° C, equivalent to a maximum energy of binding of about 50 kcal/mole while observed energies of desorption of nitrogen from tungsten all lie above about 80 kcal/mole (Ehrlich 1955)). For these reasons it is felt that this latter mechanism is unacceptable.

In the case of hydrogen, since pumping was evident in the absence of electrode potentials, and since the pumping speed increased rapidly with filament temperature, it is patent that pumping is the result of dissociation of hydrogen molecules at the filament with subsequent adsorption of atomic hydrogen at the walls, as has been established in earlier work. The temperature of the filament was not measured accurately in this work and so the observed manner of variation of pumping speed with temperature is not reported here. However, no pumping was observed below a temperature of about 1100° C, but above this temperature pumping increased rapidly up to a speed of about $\frac{1}{10}$ l./s at about 1600° C. There was apparently no dependence of speed upon pressure, but the speed decreased with an increasing gas charge on the walls, both of which are to be expected if the pumping mechanism is due to the dissociation mechanism and available sites on the walls fill with increasing atom bombardment.

The recovery of gas after sorption indicates a range of energies of binding up to about 30 kcal/mole for the hydrogen adatoms, while Hickmott's (1960) results indicate values of 25 kcal/mole and above for this energy, in reasonable agreement with the present estimate. As indicated by Hickmott the pump-down characteristics in hydrogen will thus be the result of competition between the dissociation at the filament and the recombination of hydrogen adatoms and atoms striking them from the gas phase. The speed of the former process, however, renders observations in a static system almost impossible.

Optimum pumping conditions

It has been shown that in a static pump-down, in both the rare gases and nitrogen, the ultimate pressure attainable is determined by thermal desorption of gas trapped in low energy sites (<25 kcal/mole). Two methods of reducing the ultimate pressure were tried and found successful. For a given quantity sorbed in low energy sites, the lowest pressure will be obtained with a high bombardment rate (and consequently sticking rate). By increasing the electron emission rate it was found that the ultimate pressure attainable could be substantially reduced, in a shorter pumping time of course. A similar result was obtained by operating the gauge at higher than normal electron collector potential with the consequent increase in sticking efficiency. The second method relied on the fact that if fewer low energy sites were filled the thermal desorption rate would be decreased, thus allowing a lower ultimate pressure. To achieve this, the initial gauge pumping in argon was carried out at a tube temperature of about 125° C so that particles entering low energy sites would be quickly released and all the sorption would be into comparatively high energy sites. When a fairly low pressure had been reached, i.e. most of the gas had been sorbed, the gauge temperature was reduced and the pumping continued at a faster rate because of the increased site availability. The desorption rate was now much lower because of the decreased low energy site population and in the same total time taken as for a pump-down entirely at room temperature the pressure attained was considerably reduced (Fig. 6 shows these effects). If the temperature during the initial pumping is increased to 200° C, the time taken to reach a similar ultimate pressure is increased because of the reduction in initial pumping speed; however, operation at higher emission currents would overcome this difficulty. It thus appears that better performance of the gauge can be obtained by initial operation at high emission, high electron energy and higher temperature.

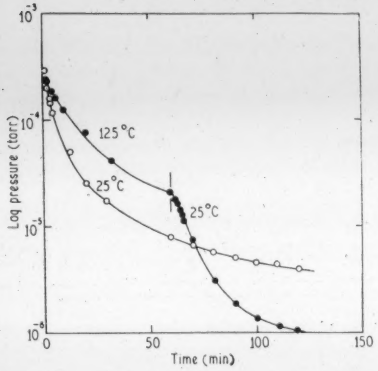


Fig. 6. The pump-down in nitrogen when the gauge was continuously operated at room temperature and when the gauge was initially operated at 125° C and subsequently reduced to room temperature.

Conclusions

The pumping action of the Bayard–Alpert gauge for the rare gases, nitrogen and hydrogen has been studied under various conditions of discharge parameters and tube temperature. The results indicate that the rare gases pump because of collision of gas ions with the walls, but no suitable mechanism can be suggested for nitrogen. Dissociation of hydrogen at the tungsten filament and adsorption of the gas atoms at the glass accounts for pumping in this gas. The results indicate that a number of kinetic processes, including sorption, gas sputtering and thermal evolution, contribute to the pumping action and their relative importance was discussed. It would appear that in a mixture of gases the overall pumping action would be extremely complex. Optimum pumping conditions for the gauge were studied and it was found that better operation was achieved by initially pumping at a high tube temperature.

Appendix

The effect of physical adsorption at refrigerated walls on pumping speed measurements

Before a pump-down was commenced in the static system the test gas was admitted to the required starting pressure with the gauge inactive. During this admission period some gas would physically adsorb upon the walls of the refrigerated trap until a layer in dynamic equilibrium with the obtaining gas pressure was established. On subsequent operation of the ion pump gas would be removed from the system with a consequent reduction in pressure. Simultaneous evolution from the trap would therefore occur since the walls would have become overconcentrated. The observed pumping rate would be the resultant of the simultaneous gauge pumping and the trap evolution.

Since the adsorption isotherm for nitrogen on glass (Langmuir 1918) is of a form which tends to saturation at low pressures ($\sim 10^{-4}$ torr) at $-196°$ K, then the above effect is appreciable under the operating experimental conditions. At the lowest pressures ($\sim 10^{-5}$ torr) the evolution rate from the trap may be as high as 90% of the true pumping rate, leading to an apparent pumping speed of only 10% of the true value. At higher pressures ($\sim 10^{-4}$ torr) the equilibrium concentration of gas upon the walls becomes increasingly less dependent upon the gas pressure, and the evolution rate decreases in comparison with the gauge pumping rate. The observed pumping speed therefore increases towards the true pumping speed. The form of curve 2 of Fig. 5(*a*) is explained on this basis.

If the trap is refrigerated only to $-78°$ C, physical adsorption on the walls is greatly diminished (saturation pressures are six orders of magnitude higher than at $-196°$ C) and true values of ion pumping speed can be deduced (curve 1 of Fig. 5(*a*)). Gases such as argon, neon and helium, which have low energies of physical adsorption, exhibit these effects only at temperatures appreciably below $-196°$ C. An estimate of the adsorption energy of nitrogen on glass was made from the observed pressure reduction on cooling a known area of glass, at about 3·5 kcal/mole, which is not an unrealistic value for the adsorption energy of a gas on glass (de Boer 1953).

Hobson and Redhead (1960) have also reported complex pumping effects when ion and adsorption pumping co-operate, but were unable to account fully for the apparent inter-relationship between the two. The above explanation, it is believed, assists in a clearer understanding of these effects.

Acknowledgments

The authors are grateful to Professor J. D. Craggs for providing facilities for this work and for much advice and encouragement. One of us (B. C.) acknowledges the receipt of a financial grant from the Institute of Nuclear Science, Belgrade.

References

Alpert, D., 1953, *J. Appl. Phys.*, **24**, 860.
Bayard, R. T., and Alpert, D., 1950, *Rev. Sci. Instrum.*, **21**, 571.
Bills, D. G., and Carleton, N. P., 1950, *J. Appl. Phys.*, **29**, 571.
Bloomer, R. N., and Haine, M. E., 1953, *Vacuum*, **3**, 128.
de Boer, J. H., 1953, *The Dynamical Character of Adsorption* (Oxford: Clarendon Press).
Carmichael, J. H., and Trendelenberg, E. A., 1958, *J. Appl. Phys.*, **29**, 1570.
Carter, G., and Leck, J. H., 1961, *Amer. Vac. Society Symposium* 1960, in the press.
Cobic, B., Carter, G., and Leck, J. H., 1961, *Brit. J. Appl. Phys.*, **38**, 288
Ehrlich, G., 1955, *J. Chem. Phys.*, **23**, 1543.
—— 1961, *J. Chem. Phys.*, **34**, 29.
Hickmott, T. W., 1960, *J. Appl. Phys.*, **31**, 128.
Hobson, J. P., and Redhead, P. A., 1960, *Advances in Vacuum Science and Technology*, Vol. I (London: Pergamon Press).
Jenkins, R. O., and Trodden, W. G., 1960, *Vacuum*, **10**, 319.
Langmuir, I., 1913, *J. Amer. Chem. Soc.*, **35**, 931.
—— 1915, *J. Amer. Chem. Soc.*, **37**, 417.
—— 1918, *J. Amer. Chem. Soc.*, **40**, 1361.
Leck, J. H., 1957, *Pressure Measurement in Vacuum Systems* (London: The Institute of Physics).
Massey, H. S. W., 1956, *Handbuch der Physik*, **36**, 307 (Berlin: Springer-Verlag).
Norton, F. J., 1954, *Committee on Vacuum Techniques. Symposium Transactions*, **47.**
Smith, P. T., 1930, *Phys. Rev.*, **36**, 1293.
Varnerin, L. J., and Carmichael, J. H., 1955, *J. Appl. Phys.*, **26**, 782.
Young, J. R., 1956, *J. Appl. Phys.*, **27**, 926.

Various electrical analogues, incorporating negative resistances, for the solution of problems in elasticity

by S. C. REDSHAW, D.Sc., and K. R. RUSHTON, Ph.D., Department of Civil Engineering, University of Birmingham

MS. received 28th February 1961

Abstract

Methods of solving many well-known problems in elasticity by means of passive resistance networks are reviewed, and the difficulties encountered when the governing partial differential equation is of the hyperbolic type are discussed.

The use of a new and very simple negative resistance, which has been incorporated into the pure resistance networks, is described and the results of several representative experiments are given.

Introduction

MANY problems in elasticity have been solved by electrical analogic methods using pure passive resistance networks and well-known examples of the type of problem which can be solved by this method are those of the St. Venant torsion problem (Redshaw 1960), the deflection and extension of flat plates (Palmer and Redshaw 1955) and the deflection of beams on elastic foundations (Palmer 1955). Although many problems of engineering importance can be solved readily with a pure resistance network, nevertheless there are many additional problems which are not amenable to solution by this method. For example, problems of the extension of a plate with known boundary tractions, or the flexure of a plate with simply supported or built in edges present no difficulty in solution by the pure resistance network; but the extension of a plate with boundary displacements specified, or the flexure of a plate with free edges, are not readily representated by this analogue (Redshaw and Rushton 1961).

In addition, there are certain problems such as that of the laterally loaded strut and the solution of eigenvalue problems concerning the critical buckling loads of struts, which cannot be attempted by the use of the positive resistance network. The difficulty arises because of the form of the governing differential equation. In the torsion problem, for example, the equation is of the elliptic type, being either that of Laplace or Poisson; whereas in the case of the laterally loaded strut an equation of the hyperbolic type holds. Again, in the extension and flexure of plates, a solution can normally be obtained either by solving the biharmonic or Lagrange form of the fourth order partial differential equation or, alternatively, two second order simultaneous partial differential equations can be solved. In electrical analogic methods there are four main methods of attack:

(i) The fourth order equation can be split into two second order equations and a pure resistance cascaded network used.

(ii) The governing equation may be solved immediately providing an inductance–capacitance net is devised.

(iii) A cascaded network may be used; the network may include resistors, capacitors and inductors, and would be coupled not by resistances but by essentially perfect transformers.

(iv) The equations can be solved by means of pure resistance networks interconnected by positive and negative resistances.

One feature which all these methods have in common is that the equations which are solved are the finite difference approximations of the actual partial differential equations.

As has been mentioned, the first method can only be used for a certain type of problem. The use of an inductance–capacitance network is well known in theory (Kron 1944) but, except for a few isolated examples, where a very coarse network has been employed (Carter 1944), no progress appears to have been made with this method because of the very great difficulty which arises when inductors and capacitors have to be matched to very close limits together with experimental difficulty of coping with unwanted parasitic electrical effects.

The third method, which centres round the use of a perfect transformer, has been used in a number of examples by several investigators (Russell and MacNeal 1953). Here the great difficulty is in obtaining the large number of essentially perfect transformers which are required if anything other than a very coarse net is needed and it would appear, from reported work, that it is necessary to feed in correcting currents to allow for the departure from the ideal of the transformers. Also, as the networks are not purely resistive, the difficulty of out of phase effects is present.

The fourth method is attractive because pure resistance networks can be used provided a suitable negative resistance can be found. The negative resistance is, in effect, a device for causing a current to flow in a direction opposite to that in which it would flow in a positive resistance of equal value. Several different types of negative resistance have been devised for various purposes but it would appear that the only one which has been used for problems in elasticity is that described by Swenson (1952) but, for reasons which will be described later, there are great practical difficulties in using this method.

The transistor negative resistance which was used in the fourth method, which is described in this paper, was devised in the Electrical Engineering Department of the University of Birmingham and it provides a simple, cheap and accurate representation of the required element.

2. The transistor negative resistance

With direct current circuits, the characteristic of a negative resistance can be obtained by several different methods. Swenson's circuit consisted of adjustable resistances and batteries which, when adjusted correctly, behaved as a

negative resistance. However, if several of these units occur in one circuit the setting of one upsets the adjustment of others.

Recently, Indiresan (1959) has devised a negative resistance which depends on stabilized positive feedback. By utilizing miniature dry batteries a single unit consists of only two transistors, eight pure resistances and one battery; a circuit diagram of a single unit is given in Fig. 1. It is interesting

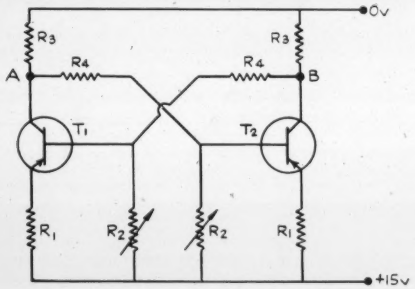


Fig. 1. Transistor negative resistance of -2000 ohms.
Negative resistance characteristic A and B. T_1 and T_2 matched pairs of O.C. 72 transistors. $R_1 = 100\,\Omega$, $R_2 = 670\,\Omega$ (variable), $R_3 = 2000\,\Omega$, $R_4 = 1000\,\Omega$.

to note that the components involved lend themselves readily to miniaturization. The unit has the following assets:

(i) It will operate with direct current.

(ii) It is independent of external connections and is, therefore, unaffected by any other similar unit in the same network.

(iii) It is compact, economical in cost and power consumption and it can, therefore, be used readily in large numbers.

From the voltage–current characteristic (Fig. 2) it should be noted that the relationship ceases to be linear above a

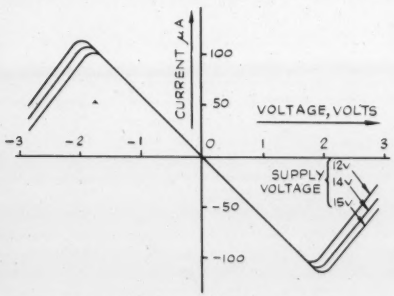


Fig. 2. Negative resistance characteristic.

certain maximum voltage. Since it is difficult to match all the components exactly, a small current does flow with zero voltage across the terminals; however, this error can be cancelled out by changing the sign of the applied voltage once every second. The effective tolerance of the negative resistance, which is less than $\pm 1\%$, is of the same magnitude as that of the positive resistances used in the circuits.

If the value of the normal positive resistor is 100 ohms, the value actually used in the networks, the resistance value of the negative resistor would be in the region of 1000 ohms.

3. The negative resistance applied to strut problems

The use and potentialities of the negative resistance can be best demonstrated by giving a description of some typical experiments and, in the following sections, the problem of the laterally loaded strut and the determination of the critical buckling load of a strut will be described.

3.1. *Laterally loaded strut*

A pure resistance analogue for the flexure of beams has previously been described by Palmer (1955). Whilst investigating the axial loading of beams, he found that a tensile load could be represented automatically, but a compressive load demanded an iterative procedure. However, with the aid of a negative resistance the compressive axial load can be represented automatically.

The differential equation governing the deflection of a beam under an axial load P, a distributed lateral load q and end moment M_E is

$$EI\frac{d^4w}{dx^4} - P\frac{d^2w}{dx^2} - q = 0 \tag{1}$$

where P is positive when tensile.

Since
$$M = -EI\frac{d^2w}{dx^2} \tag{2}$$

Eqn (1) can be re-written as

$$\frac{d^2M}{dx^2} - \frac{P}{EI}M + q = 0. \tag{3}$$

Eqns (2) and (3) in finite difference form become

$$M_1 + M_3 - 2M_0 - \frac{Ph^2}{EI}M_0 + qh^2 = 0 \tag{4}$$

$$w_1 + w_3 - 2w_0 + \frac{h^2}{EI}M_0 = 0. \tag{5}$$

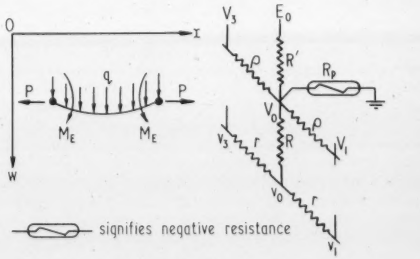


Fig. 3. Electrical analogy for laterally loaded strut.

Writing Kirchhoff's law for the resistance network in Fig. 3, we have

$$V_1 + V_3 - 2V_0 - \frac{\rho}{R_P}V_0 + \frac{\rho}{R'}E_0 = 0 \tag{6}$$

$$v_1 + v_3 - 2v_0 + \frac{r}{R}V_0 = 0. \tag{7}$$

Equations (4) and (6), also (5) and (7), are seen to be analogous and the following identities can be noted:

$$v = w \qquad V = \frac{R}{r}\frac{h^2}{EI}M$$

$$E = \frac{R'R}{\rho r}h^4\frac{q}{EI} \qquad R_P = \frac{\rho}{r}\frac{R}{P}.$$

If the axial load P is compressive, that is if it is negative, then the resistance R_P must have a negative value. A typical experiment is specified in Fig. 4(*a*). The operation of the

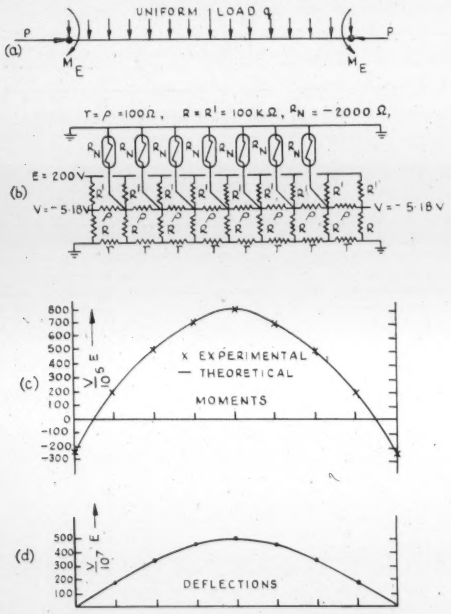


Fig. 4. Laterally loaded strut, result of experiment.

analogue to analyse the problem of a laterally loaded strut is particularly simple since it only entails the setting up of certain potentials to the network, no iteration being necessary. To the network shown in Fig. 4(*b*) a potential is applied to all nodes of the upper net to represent the uniform lateral load and the earthed negative resistances simulate the axial load. The end moments are applied as voltages to the corresponding points on the middle net and the earthing of the boundaries of the lower net is analogous to the pinned ends.

Once the potentials have been applied to the boundaries the voltages of the internal nodes are directly proportional to the moments and deflections. The potentials are most easily measured using a null method with a potential dividing resistance.

For the electrical analogue solution it will be seen that the beam was divided into eight mesh lengths and a comparison between the analogue results and the computation of the moments and deflections derived from a Howard (1928) polar diagram are given in Figs 4(*c*) and 4(*d*).

The agreement between theoretical and experimental values is good, the maximum error being less than 1%.

Although the case of a strut of uniform flexural rigidity has been examined in this example, nevertheless there is no restriction in applying the method to a strut of non-uniform section; all that would be involved is a modification of the various resistance values to take account of the variable strut section.

3.2. *Critical buckling load of strut*

By a simple extension of the analogy it is possible to determine the critical buckling load of a strut. A network identical to that used in the previous example is required, the axial loading being increased in increments by a change in value of the corresponding resistance. The increasing load results in larger deflections and hence larger moments; the variation in the resistance value of the resistors, within the tolerance of ±1%, is analogous to the small physical imperfections of the strut and ensures that deflections do occur as soon as an axial load is applied. In operating the analogue a measure of the strut instability is needed and, rather than record deflection, it is more convenient to record the increase in moment at the centre of the strut. Fig. 5,

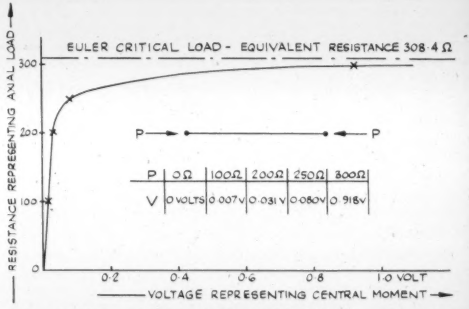


Fig. 5. Euler critical load of strut determined by electrical analogy.

which relates the voltage corresponding to the central moment with the resistance which is analogous to the applied load, shows the same characteristics as a normal critical load diagram. The Euler critical load, which was calculated and transformed into terms of resistance, can be seen to be the load at which the moments would become infinite, resulting in failure of the strut. It is of interest to note that when resistances are incorporated which correspond to the Euler load, the circuit becomes unstable.

It should be noted that all the required information can be obtained from the analogy for determining the critical load by means of a Southwell (1932) plot, if required; this is advantageous as the experiment need not then proceed to an unstable state which might be injurious to the negative resistances.

4. The negative resistance applied to the extension and flexure of flat plates

There are two alternative ways of describing mathematically the extension and flexure of flat plates and these have been discussed in detail by Southwell (1950). The fourth order partial differential equation has been solved by Palmer and

Redshaw (1955) using a pure resistance cascaded network but the alternative expression of two simultaneous second order partial differential equations require negative resistances before they can be solved by means of an electrical analogue. This second representation has the advantage that the extensional problem with a boundary displacement specified, and the flexural problem with a free boundary, can easily be solved.

The two simultaneous equations for the flexural problem are

$$\frac{\partial}{\partial x}\left(\frac{\partial U}{\partial x}+\frac{\partial V}{\partial y}\right)+\left(\frac{1+\nu}{1-\nu}\right)\nabla^2 U + \left(\frac{2}{1-\nu}\right)\frac{\partial}{\partial x}(\nu\Omega_1-\Omega_2)=0 \quad (8)$$

$$\frac{\partial}{\partial y}\left(\frac{\partial U}{\partial x}+\frac{\partial V}{\partial y}\right)+\left(\frac{1+\nu}{1-\nu}\right)\nabla^2 V + \left(\frac{2}{1-\nu}\right)\frac{\partial}{\partial y}(\nu\Omega_2-\Omega_1)=0 \quad (9)$$

where U and V are functions of the moments, Ω_1 and Ω_2 depending on the loading of the plate.

By the application of Kirchhoff's laws, it will be seen that the electrical network illustrated in Fig. 6 is identical in form to the finite difference form of Eqns (8) and (9), thus:

$$U_1+U_3+\frac{1+\nu}{2}(U_2+U_4)-(3+\nu)U_0 + \frac{1-\nu}{8}(V_5-V_6+V_7-V_8)+h^2\left\{\frac{\partial}{\partial x}(\nu\Omega_1-\Omega_2)\right\}=0 \quad (10)$$

$$V_1+V_3+\frac{1+\nu}{2}(V_2+V_4)-(3+\nu)V_0 + \frac{1-\nu}{8}(U_5-U_6+U_7-U_8)+h^2\left\{\frac{\partial}{\partial y}(\nu\Omega_2-\Omega_1)\right\}=0. \quad (11)$$

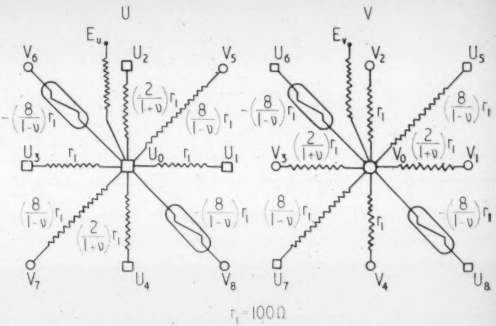
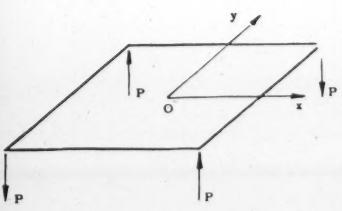
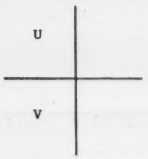


Fig. 6. Coupled network for plate problem, electrical networks for a node.

The problem used to illustrate the ability of a circuit to solve the U–V equations (Eqns (8) and (9)) is that of a square plate with concentrated loads on two opposite corners and point supports at the remaining two corners. The plate is illustrated in Fig. 7(*a*), and the boundary conditions in terms of U and V in Fig. 7(*b*). The values of U and V measured on the analogue are to be seen in Fig. 7(*c*). On all sections



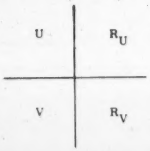
(a) DETAILS OF PROBLEM

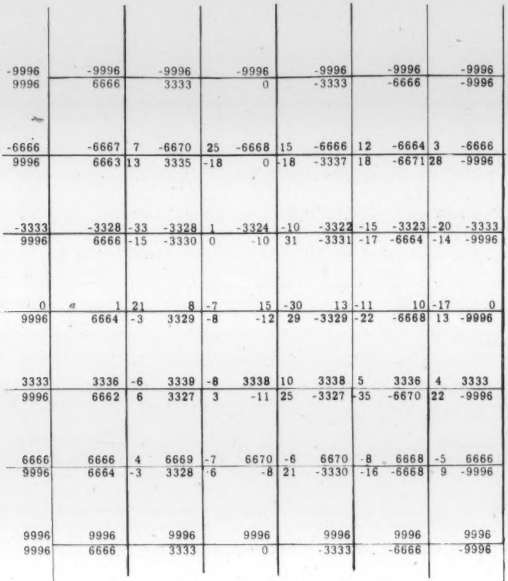


-10000 / 10000	-10000 / 6667	-10000 / 3333	-10000 / 0	-10000 / -3333	-10000 / -6667	-10000 / -10000
-6667 / 10000						-6667 / -10000
-3333 / 10000						-3333 / -10000
0 / 10000						0 / -10000
3333 / 10000						3333 / -10000
6667 / 10000						6667 / -10000
10000 / 10000	10000 / 6667	10000 / 3333	10000 / 0	10000 / -3333	10000 / -6667	10000 / -10000

(b) BOUNDARY VALUES

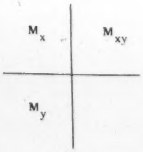
Fig. 7. Electrical analogue results for twisted square plate.



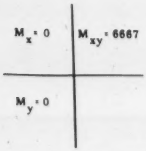


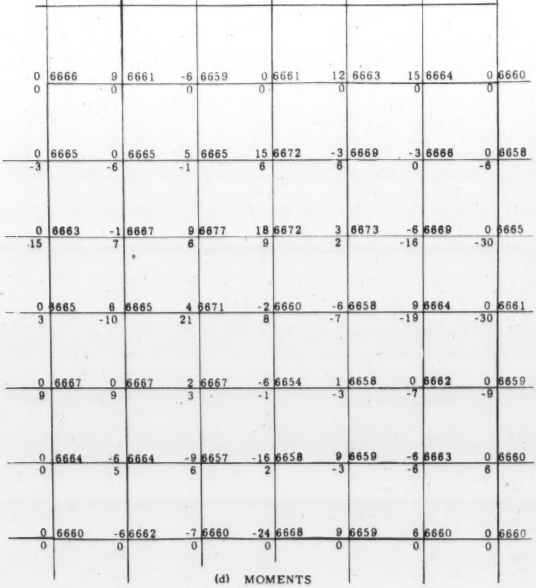
(c) MEASURED POTENTIALS and RESIDUALS

EXPERIMENTAL



THEORETICAL





(d) MOMENTS

Fig. 7. Electrical analogue result for twisted square plate.

parallel to the x-axis the values of U are constant, and parallel to the y-axis the magnitude of V is also unchanged. It can be seen that the analogue results are good. Further, the residuals R_u and R_v have been calculated and can be seen to be generally small.

By differentiating the values of U and V, moments were calculated (Fig. 7(*d*)).

The direct moments should both be zero over the whole plate and the twisting moment should be 6667 at every node.

5. Conclusions

The negative resistance which has been described in this paper is economical to manufacture and extremely simple to operate and, in fact, its use is very little more complicated than the use of a pure positive resistance. The possession of the negative resistance opens up a large field of problems which can be solved by its use. Further work is proceeding in the solution of plate problems by the U–V technique particularly with regard to the formulation of the more difficult boundary conditions.

References

CARTER, G. K., 1944, *J. Appl. Mech.*, **2**, A162.

HOWARD, H. B., 1928, *Aeronautical Research Committee R & M* No. 1233.

INDIRESAN, P. V., 1959, *J. Brit. Instn Radio Engrs*, **19**, 401.

KRON, G., 1944, *J. Appl. Mech.*, **2**, A149.

PALMER, P. J., 1955, *Solution of Elastic Foundation Problems by means of a Resistance Network* (Journées Internationales de Calcul Analogique, Bruxelles).

PALMER, P. J., and REDSHAW, S. C., 1955, *Aeronaut. Quart.*, **6**, 13.

REDSHAW, S. C., 1960, *Brit. J. Appl. Phys.*, **11**, 461.

REDSHAW, S. C., and RUSHTON, K. R., 1961, *Aeronaut. Quart.*, **12**, 275.

RUSSELL, W. T., and MACNEAL, R. H., 1953, *J. Appl. Mech.*, **20**, 349.

SOUTHWELL, R. V., 1932, *Proc. Roy. Soc. Lond., A*, **135**, 601.

—— 1950, *Quart. J. Mech.*, **3**, 257.

SWENSON, G. W., 1952, *J. Aero. Sci.*, **19**, 273.

Sorption and replacement of ionized noble gases at a tungsten surface

by R. B. BURTT, MSc., Ph.D., J. S. COLLIGON, B.Eng., and J. H. LECK, M.Eng., Ph.D., A.Inst.P., A.M.I.E.E., Department of Electrical Engineering, Liverpool University

MS. received 10*th February* 1961

Abstract

Experiments are described in which ion beams of the order of five micro-amperes with a low energy spread have been used to bombard a tungsten surface. Sorption values for the surface are given for neon, argon and krypton and show that a maximum sorption exists at any particular bombarding ion energy. Using several simplifications a value for the depth of penetration of the ions is derived from the results and this is shown to be in reasonable agreement with values quoted by other workers. Results from a second series of experiments where ions of a second gas release the sorbed ions of the first gas are also presented.

Introduction

In spite of the many recent advances that have been made in the techniques of high vacuum and the great interest shown in the practicability of ion pumping there is not yet a clear understanding of the reaction taking place when surfaces are bombarded with high energy positive ions. The experiments described in this paper are effectively a continuation of work carried out in these laboratories in which metal surfaces were bombarded by positive ions of the noble gases and complex sorption and replacement phenomena observed (Brown and Leck 1955, Leck 1956). Measurements reported here represent an improvement in technique because the energy spread of the incident ion beam (one of the chief disadvantages of the previous work) has been reduced to negligible proportions. Also the ion beam has now been confined to a small and closely defined area on a ribbon target which is carefully shielded from all extraneous bombardment. Previously the area under ion bombardment was not accurately known. With the above improvements the present apparatus is well equipped to investigate sorption effects under ion bombardment and gives consistent and unambiguous results.

Apparatus and experimental technique

For this work the target was a tungsten ribbon with a purity of 99·9% mounted under tension and having dimensions: 0·030 in. wide, 0·001 in. thick and 2 in. long. Positive ions were formed in an electrodeless glass discharge chamber comprising a Pyrex cylinder 8 in. long and 2¼ in. in diameter, closed to the discharge at the top by a fused silica plate. The radio frequency source operated at 20 Mc/s with a continuously variable power of up to 250 w and was coupled to the discharge by four turns of heavy copper wire. Optimum conditions held with a gas pressure of the order of 10^{-3} torr in the chamber. The mounting of this discharge chamber in relation to the target and to the intermediate focusing electrodes is shown in Fig. 1. These electrodes served to extract the ions, to focus them on to the target and to suppress secondary electron emission from the target. Typical operating potentials are indicated in Fig. 1; the

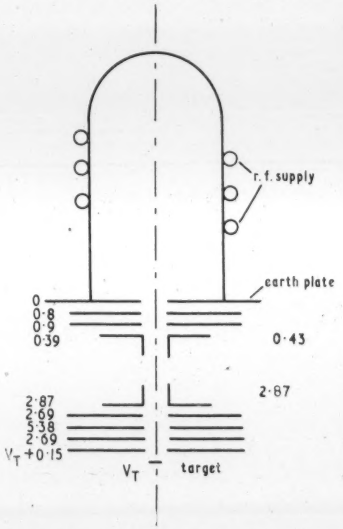


Fig. 1. Discharge chamber and ion extraction system. Typical electrode potentials are indicated on the diagram in kv negative with respect to the earthed plate.

target potential (V_T) was variable from 0 to 4 kv in order to control the energy of the bombarding ions. A rectangular cross section was chosen for the beam both to give a uniform ion distribution over the target and to minimize the space-charge spreading of the beam. The simple system for ion extraction from the discharge was based upon the work of Pierce (1940) and later of Thonemann and Harrison (1955). All the slits were manufactured from non-magnetic stainless steel. With optimum coupling between the r.f. oscillator and the discharge and a gas pressure of 10^{-3} torr a target current of 5 μA was readily obtained. Simple auxiliary experiments showed the beam to be confined to 1 cm length in the central region of the target and the energy spread to be very low (80% of the beam inside an energy band 40 v wide).

The vacuum pumping system for this assembly was, as shown in Fig. 2, of conventional design. Two small fractionating oil diffusion pumps charged with Apeizon 'B' oil were used to evacuate the system; each pump had a vapour trap associated with it which could be held at either solid carbon dioxide or liquid nitrogen temperature. The pumping line of the first diffusion pump led straight into

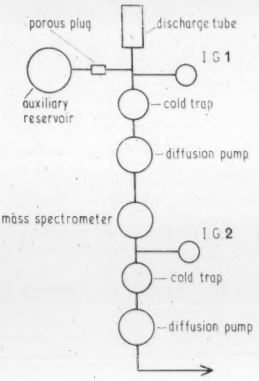


Fig. 2. The vacuum system.

the source of the mass spectrometer which was a conventional 'Metrovac M.S.3' with a Nier-type source and a 6 in. radius, 60 degree deflection analyser tube. As this was an all-metal completely demountable system with rubber sealing rings, the baking temperature was limited to 100° c for both the spectrometer and the discharge chamber. In normal operation it was observed that the pressure, as measured by an ion gauge (I.G.1), could be reduced to below 10^{-7} torr with carbon dioxide in the traps and to 10^{-8} torr with liquid nitrogen.

Gas was admitted to the top of the discharge chamber from an auxiliary reservoir through a fine leak comprising a small plug of silicon carbide sealed into a glass tube. Spectroscopically pure gas samples were used and, by having a continuous-flow vacuum system, a very pure sample could be maintained in the discharge chamber at a pressure controlled and variable through the range 10^{-4} to 10^{-2} torr (1–100 torr in the auxiliary reservoir). The gas conductance of the exit slits from the discharge chamber was such that a steady sample pressure in the discharge of 10^{-3} torr set up a pressure at the target of 10^{-5} torr.

To obtain a sorption characteristic the following method was adopted:

(i) After first baking the target at 1600° c for five minutes and ensuring that the background pressure was reasonably low (of the order of 10^{-7} torr) sample gas was admitted to the discharge chamber, the target potential adjusted to the pre-determined value and the discharge switched on.

(ii) Careful note of the target current was taken and, after a definite ion bombardment, the discharge was switched off and the sample gas pumped from the system.

(iii) When the partial pressure of the sample gas had fallen to less than 10^{-9} torr the target was heated to 1600°c by passing an alternating current through it. Any sample gas desorbed during this heating was observed as a pressure transient in the mass spectrometer. Alternatively, in another series of experiments, instead of liberating the sorbed gas by heating it was released by bombarding the target with a second gas. After the subsidiary bombardment the target was heated and desorption of the remaining first gas observed by the mass spectrometer.

Since all the desorbed gas must pass through the spectrometer the integrated pressure transient with respect to time must, after introducing a suitable constant (related to the pumping speed of the sample gas in the system), represent accurately the total gas desorbed from the target. The calibration is relatively straightforward as the flow through the porous plug can easily be determined and used as a standard.

Results

In the first series of experiments the desorption of the gases neon, argon and krypton was observed for various conditions of bombardment. In each case the gas was recovered and recorded in a single step by heating the target to 1600° c.

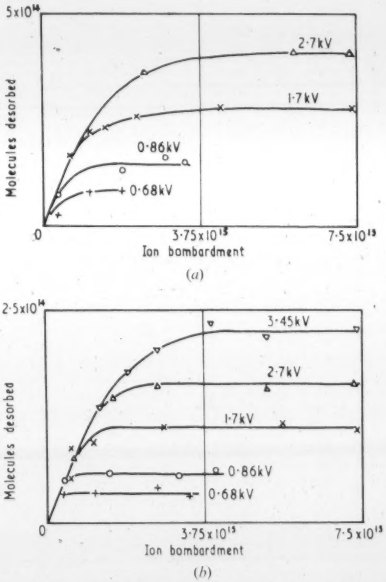


Fig. 3. The relationship between the number of molecules desorbed and the number of ions bombarding the target. The energy of the incident ions is shown on each curve. The effective target area was approximately 0·075 cm².
(*a*) argon. (*b*) krypton.

For each gas the desorption was measured as a function of both the number and also the energy of the incident ions. Results are summarized in Figs 3 and 4. Fig. 3 shows how for low values of ion bombardment an approximately linear relation between bombardment and sorption exists but at higher bombardments a saturation level is reached, its magnitude dependent on the energy of the bombarding ions. All these curves are typical of a large number of observations.

Fig. 4 summarizes the results showing the saturation sorption level plotted as a function of the ion energy for each gas.

The curves in Fig. 5 show that the desorption of argon

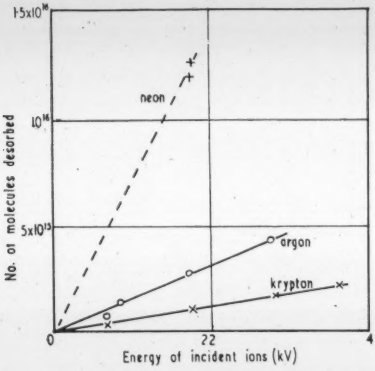


Fig. 4. The saturation sorption level shown as a function of the energy of the incident ions.

from the target took place over a comparatively large temperature range. To obtain these results the target was heated in steps, each of approximately 200 deg C, and the gas desorbed at each step recorded and plotted as a function of the absolute temperature. All the curves in Fig. 5(*a*) represent the desorption after high energy bombardment (2·7 kev in each case); they show, from (1) to (3), the effect of progressively increasing the bombardment up to saturation level. Alternatively Fig. 5(*b*) shows the effect of increasing

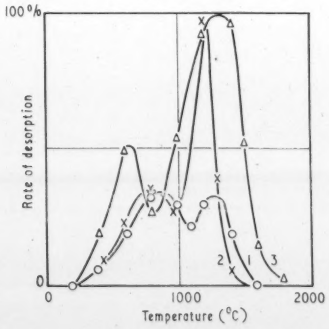


Fig. 5(*a*). The rate of desorption of argon with increasing temperature after ion bombardment at 2·7 kev:
1, bombardment to 33% saturation; 2, bombardment to 67% saturation; 3, bombardment to saturation.

the energy of ion bombardment, the intensity of the bombardment being such as to reach saturation in each case. Fig. 6 shows the desorption–temperature relations for the other gases. Neon, like argon, is seen to be liberated in two favoured temperature ranges, whilst krypton is mainly released in one temperature band.

In the second series of experiments the desorption was measured only after a further bombardment with ions of a second gas. It was found that any of the inert gases, neon argon, or krypton, would remove any other although the efficiencies were by no means equal. The effect of krypton bombardment following a primary argon sorption and the converse is shown in Fig. 7. In each case the number of

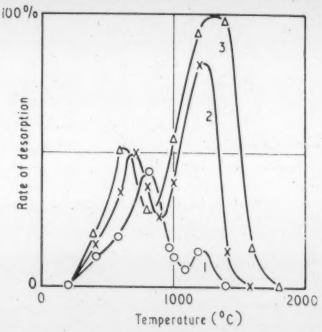


Fig. 5(*b*). The rate of desorption of argon with increasing temperature after bombardment to saturation:
ion energy: 1, 0·86 kev; 2, 1·7 kev; 3, 2·7 kev.

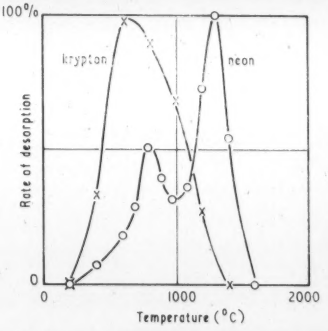


Fig. 6. The rate of desorption of neon and krypton with increasing temperature after the target has been saturated by bombardment with ions of 1·7 kev energy.

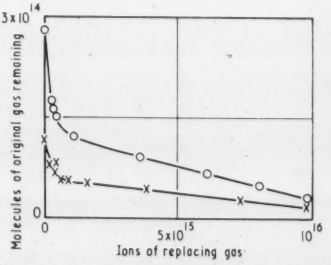


Fig. 7. The number of primary ions remaining sorbed on the target as a function of the secondary ion bombardment.
○ argon replacing krypton.
× krypton replacing argon.

molecules recovered in the final baking cycle is shown as a function of the number of ions of the second gas striking the target. The primary bombardment was naturally made the same in each case (sufficient just to reach sorption level). The condition of the desorption–temperature characteristic does not change fundamentally as a result of the secondary bombardment.

Discussion

The results that have been set out in the previous sections fit in well with an explanation of sorption under ion bombardment first suggested by Le Claire and Rowe (1955) and later considered in greater detail by, for example, Carmichael and Trendelenburg (1958) and Blodgett and Vanderslice (1960). This theory postulates that ions, because of their high kinetic energy, are driven to some considerable depth into the interior of the metal target and, at the same time, sputter away atoms from the crystal surface. Thus although in the initial stages of the bombardment the sorption rate will be comparatively high, it must progressively fall as the surface is eroded away, so exposing and releasing molecules previously sorbed. Eventually an equilibrium must be reached when the density of trapped molecules is such that the rate of release is equal to the rate of capture.

Direct evidence for dynamic equilibrium at a surface is obtained when the type of ion in the discharge is changed. After the change desorption of the original gas persists for a considerable time and can readily be detected. The curves in Fig. 7, which show that the total sorption of a particular gas is reduced by bombarding with a different gas, demonstrate this effect. This simple explanation of a sorption–desorption equilibrium can only hold for conditions such as those existing in these experiments where the target surface is only bombarded by the primary ion beam. For example the characteristics are changed considerably when the surface is also receiving neutral particles sputtered from other electrodes (Le Claire and Rowe 1955, Blodgett and Vanderslice 1960).

For the present experiments the equilibrium (or saturation) level will clearly increase with increasing energy of the incident ions since with a greater energy a greater penetration is to be expected. However, the effect will be counterbalanced to a certain extent by the increased surface erosion rate. Although there is not sufficient sputtering data available for tungsten at the present time to allow the erosion rate to be calculated precisely an estimate to an accuracy of about a factor of two can be made. From the data available it is reasonable to assume a sputtering factor S of 2·0 for both argon and krypton at an energy of about 2·5 kev. Thus the rate of surface erosion is given quite simply by

$$r = n_+ S/U \text{ cm/s} \qquad (1)$$

where n_+ is the rate of arrival of ions per unit area at the surface and U is the number of tungsten atoms per unit volume in the crystal. With a knowledge of r and provided a reasonable estimate of the ion penetration depth can be made the build up of the sorbed layer to saturation, and the saturation level, can be calculated.

In order to carry out this calculation in relation to the above results the penetration depth distribution curve shown in Fig. 8 has been assumed, i.e. the probability is a maximum at zero penetration and falls off linearly to zero at some arbitrary depth. This simple postulate is not unreasonable in view of the work of Young (1956) and Bradov and Okuneva (1955). It is a straightforward calculation to show that the saturation sorption level is reached after the surface has eroded a depth λ given by

$$\lambda = S/Un_+ \qquad (2)$$

and is approached in the manner shown in Fig. 9. For comparison the experimental as well as the theoretical curves are plotted in Fig. 9. For convenience the scales are

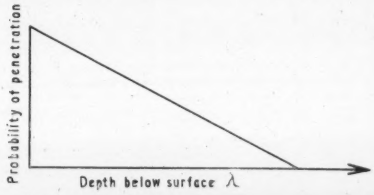


Fig. 8. Assumed probability curve for the penetration of ions into the surface layers of tungsten.

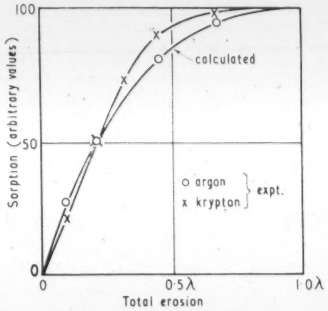


Fig. 9. A comparison of the calculated and measured values of sorption level as a function of the surface erosion.

adjusted so that all curves agree at 50% saturation level. The two experimental curves shown here have been chosen deliberately as they represent about the best and the worst fit to the theoretical values of all the results obtained. The value of λ calculated from Eqn (2) is $1{\cdot}25 \times 10^{-6}$ (at 2·7 kev) for both argon and krypton.*

Using these calculated values of λ the relation between the gas remaining and the total bombardment when the nature of the bombarding ion is changed (i.e. corresponding to the experimental data of Fig. 7) has also been calculated and plotted alongside the experimental results in Fig. 10. There is a very satisfactory agreement between the calculated and the experimental curves to an erosion depth of approximately $0{\cdot}25\,\lambda$. Beyond this point the deviation increases rapidly suggesting that a small fraction of the incident ions penetrates much further than the estimated λ.

The values of penetration depth emerging from the above comparison of theory with practice can be compared with the values measured by other workers using entirely different

* In fact Eqn (2) was modified to $0{\cdot}21\,\lambda = S/Un_+$ where n_+ in this case represents the number of ions required to reach 50% of maximum sorption and $U = 6 \times 10^{22}/\text{cm}^3$.

techniques. This comparison is summarized briefly in the table. In view of the fact that ion bombardment by the various workers has been carried out at various arbitrary ion energies ranging from 150 to 4000 ev all the results have

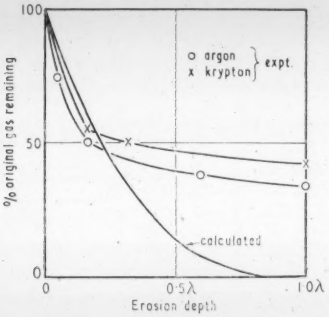


Fig. 10. A comparison of the calculated and measured values of the gas remaining sorbed as the surface is eroded by bombardment with a different gas.

been reduced to a common energy of 1000 ev by assuming a linear relation between penetration and energy. To obtain his results Young (1955, 1956) measured the penetration of the ions through thin foils whilst all the other workers used

Investigators	*Material bombarded*	*Bombarding ions*	*Depth* λ ($\times 10^6$) (*cm*)
Young (1956)	Aluminium	Hydrogen (1·0 kev)	1·5
Young (1955)	Phosphor	Argon (20 kev)	0·25
Bartholomew and La Padula (1960)	Nickel	Radioactive krypton (0·15 kev)	260
Bredov and Okuneva (1955)	Germanium	Radioactive caesium (4 kev)	2·5
Present work	Tungsten	Argon and krypton (2·7 kev)	0·63

A summary of the published values for the penetration of positive ions (in the energy range 0·1 to 20 kev) into solid materials. In order to assist the comparison the penetration depths λ given in this table have all been referred to an ion energy of 1·0 kev. It has been arbitrarily assumed that the depth of penetration is proportional to incident energy.

radioactive tracers with chemical or mechanical etching after bombardment to remove the surface layers. It is clear that the present results are in general agreement with those quoted by Young (1955, 1956) and Bredov and Okuneva (1955) and it is clear that the results published by Bartholomew and La Padula (1960) are out of step with the majority.

Curves in Fig. 5 showing the desorption as a function of the target temperature indicate that the sorbed atoms are distributed in sites of varying binding energy ranging from 35 to 100 kcal/mole. It is interesting to note that in the case of neon and argon the sites are grouped around two levels with a relatively sparsely filled band in between. The band with the lower energy range is filled first as the sorption builds up to the saturation level. This unusual 'double-peaked' pattern (which suggests two different mechanisms of binding, for example, interstitial and vacancy) was confirmed over many measurements and it is of interest to note that this kind of behaviour has been observed by other workers in sorption experiments (C. W. Tucker, Jr, and F. J. Norton, private communication 1960).

Conclusions

Care has been taken in the design of these experiments to ensure relatively simple reactions at the test surface, i.e. the bombardment was only by positive ions all contained in a fairly narrow energy range. The results obtained indicate that under these conditions sorption on tungsten reaches a very clearly defined maximum. The whole of the analysis does, however, rest on the assumption that all the gas taken up is recovered in the subsequent heating cycle. Although no direct proof of this can be offered (it is obviously difficult to isolate the sorption to the small target from that to all the other electrodes, whereas the desorption can easily be measured by selective heating) indirect evidence is in its favour. For example, earlier work showed ion bombardment to give the same total desorption, and measurements upon glass surfaces showed the desorption to be equal exactly to the previous sorption. The conditions in the above work are different from those experiments in which full desorption has not been observed. Blodgett and Vanderslice (1960), for example, literally buried the sorbed gas under many layers of metal laid down by sputtering.

The calculations—based upon the theory of dynamic equilibrium between sorption and desorption—showed better than order of magnitude agreement with other workers for the depth of penetration of incident particles. There does not, however, appear to be any simple explanation for the deviation of theory and practice in the replacement experiments (Fig. 10). There is appreciable thermal desorption even after an extremely long ion bombardment. This cannot be explained by any simple experimental errors—such as non-linearity of target bombardment—as these would all make the saturation curves in Fig. 3 less decisive.

Acknowledgments

We are grateful to Professor J. D. Craggs of this University for advice in connection with the publication of this paper, also to Mr. N. Collins for invaluable assistance with the glass-blowing.

One of us (J. S. C.) is indebted to the Department of Scientific and Industrial Research for financial assistance during the period covered by this work. Likewise, R. B. B. is grateful to both the Department of Scientific and Industrial Research and the Admiralty.

References

Bartholomew, C. Y., and La Padula, A. R., 1960, *J. Appl. Phys.*, **31**, 445.

Blodgett, K. B., and Vanderslice, T. A., 1960, *J. Appl. Phys.*, **31**, 1017.

Bradov, M. M., and Okuneva, N. M., 1955, *Dok Akad. Nauk SSSR*, 113.

Brown, E., and Leck, J. H., 1955, *Brit. J. Appl. Phys.*, **6**, 161.

Carmichael, J. H., and Trendelenburg, E. A., 1958, *J. Appl. Phys.*, **29**, 1570.

Le Claire, A. D., and Rowe, A. H., 1955, *Rev. Métall.*, **52**, 94.

Leck, J. H., 1956, *Proceedings Symposium on Chemisorption.* Univ. Coll. N. Staffs., **162.**

Pierce, J. R., 1940, *J. Appl. Phys.*, **11**, 548.

Thonemann, P. C., and Harrison, E. R. H., 1955, *A.E.R.E. Report GP/R* 1190.

Young, J. R., 1955, *J. Appl. Phys.*, **26**, 1302.

—— 1956, *J. Appl. Phys.*, **27**, 1.

Emissivity errors of infra-red pyrometers in relation to spectral response

by P. M. REYNOLDS, B.Sc., Ph.D., The British Non-Ferrous Metals Research Association, London, N.W.1

MS. received 20*th March*, 1961

Abstract

The emissivity errors of total, partial and monochromatic radiation pyrometers for use in the temperature range 500–800° K *are discussed theoretically. The pyrometer output is expressed as a power law of absolute temperature with an exponent n which varies with temperature and the spectral response of the pyrometer. The emissivity error decreases as n increases. It is shown that n increases as the spectral response of the pyrometer is restricted to shorter wavelengths in the near infra-red region.*

1. Introduction

A RADIATION pyrometer measures the energy flux emitted from a hot body, within a range of wavelengths defined by its 'optical' system and detector. This energy is only uniquely related to temperature for the special case of a black-body emitter; when real emitters are considered an additional variable, the emissivity, is introduced to allow for the departure from black-body conditions. The radiation pyrometer then indicates an apparent temperature which differs from the true temperature of the specimen. The error in the indicated temperature depends not only upon the emissivity, but also upon the wavelength response of the pyrometer. For example, if a disappearing filament pyrometer (working at a wavelength of 0·65 μ) and a total radiation pyrometer, each calibrated against a black body, both view a surface at 1000° C which has a constant spectral emissivity 0·50, the indicated temperatures are 953° C and 805° C respectively.

The errors of brightness and two-colour spectral radiation pyrometers have been discussed by Pyatt (1954) and by Emslie and Blau (1959), principally for the temperature range above 1000° K. Gill (1957) has dealt with the general problem of using a total radiation pyrometer to measure temperatures below 700° K, including the emissivity error.

Recent developments in photoelectric detectors of infra-red radiation have enabled the range of temperatures measurable by radiation pyrometers to be extended down to about 150° C (Avery, Goodwin and Rennie 1957, Harmer and Watts 1955). If these detectors are combined with infra-red filters it becomes possible to design new radiation pyrometers with a variety of different wavelength responses, and consequently different emissivity errors. Experiments show (Harmer and Watts 1955, Larson and Shenk 1940) that within the temperature range 200–1000° C a pyrometer with a lead sulphide cell detector, which is insensitive to wavelengths greater than 3·5 μ, has a smaller emissivity error than a total radiation pyrometer. This occurs because the signal output varies as a high power (8th–12th) of absolute temperature. The present paper examines the effect of the spectral characteristics of the detector system more fully, with particular reference to the problem of measuring the temperature of materials, such as aluminium, whose emissivity is low and varies markedly with surface condition, temperature and wavelength, in the temperature range 500–800° K. Alternative methods of obtaining a high power signal output law are considered.

2. Emissivity errors arising from total, partial and monochromatic radiation pyrometers

Total radiation pyrometer

If the energy emitted by the detector forms a negligible part of the total energy exchange, the net energy collected by a total radiation pyrometer from a black-body source at absolute temperature T is given by aT^4, where a is a constant. For a non-black body of total emissivity ϵ this becomes ϵaT^4. Thus if a pyrometer which has been calibrated against a black body is used to measure the temperature of a non-black body, it will indicate an apparent temperature T_A where

$$T_A{}^4 = \epsilon T^4 \tag{1}$$

and the fractional error in temperature measurement is

$$(T - T_A)/T = 1 - \epsilon^{\frac{1}{4}} \tag{2}$$

i.e. the error depends upon the fourth root of the total emissivity.

Partial radiation pyrometers

The energy $E_\lambda d\lambda$ emitted by a black body in the wavelength range λ to $\lambda + d\lambda$ is given by

$$E_\lambda d\lambda = c_1\lambda^{-5}\{\exp(c_2/\lambda T) - 1\}^{-1}d\lambda \tag{3}$$

where c_1 and c_2 are constants. Fig. 1 shows the variation of E_λ with wavelength for temperatures from 500 to 800° K.

If a photoelectric infra-red detector has spectral response R_λ at wavelength λ, the output signal is given by

$$S = b\int_0^\infty R_\lambda E_\lambda d\lambda \tag{4}$$

(b = constant), provided that the signal output is linearly related to the energy detected.

The spectral response of a Mullard 615V lead sulphide photoconductive cell is shown in Fig. 2. Harmer and Watts (1955) have shown experimentally that with this type of detector

$$S = cT^n \tag{5}$$

where c is a constant, and n, which varies with temperature, lies between 8 and 12. The error in temperature measurement is then

$$(T - T_A)/T = 1 - \epsilon^{1/n} \tag{6}$$

where ϵ is now the mean spectral emissivity over the range of wavelengths collected by the cell. Since $\epsilon < 1$, $\epsilon^{1/n}$ approaches unity as n increases, and a lead sulphide cell

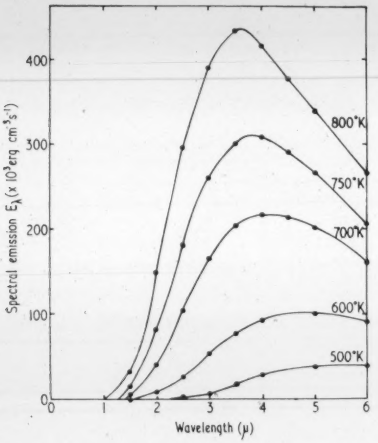


Fig. 1. Variation of energy emission with wavelength.

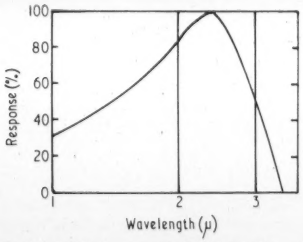


Fig. 2. Spectral response of lead sulphide cell.

pyrometer is less affected by emissivity than a total radiation pyrometer because $n > 4$.

The high value of the exponent n for the lead sulphide cell is mainly due to the cut-off in its wavelength response at 3·0–3·5 μ, and arises because the energy emitted by a black body on the short wavelength side of its maximum emission varies more rapidly with temperature than the total emission. High values of n could also be obtained with other detectors used with suitable filters.

The relationship between the signal output S of any given detector system and the absolute temperature T, can be deduced graphically by multiplying corresponding ordinates of the emission curve (Fig. 1), and the known filter transmission or spectral response curve, such as Fig. 2, and determining the area beneath the derived curve for different temperatures.

If a filter of spectral transmission F_λ, at wavelength λ, is interposed between a total radiation receiver and a black-body source, the signal output is given by

$$S = d\int_0^\infty F_\lambda E_\lambda d\lambda \qquad (7)$$

where d is a constant. Since Eqns (4) and (7) are mathematically identical, a filtered total radiation detector or a photocell detector should give the same output law, if the normalized curves for F_λ and R_λ are identical.

To show how the output law varies with F_λ(or R_λ), suppose for simplicity, that F_λ is constant to a wavelength λ_{max} and zero when $\lambda > \lambda_{max}$ (Fig. 3). The variation of

$$\log_{10}\int_0^\infty F_\lambda E_\lambda d\lambda \left(\text{or } \log_{10}\int_0^\infty R_\lambda E_\lambda d\lambda\right)$$

with $\log_{10} T$ for different values of λ_{max} is shown in Fig. 4. The slope of these curves is equal to n; approximate values of n are given in Table 1.

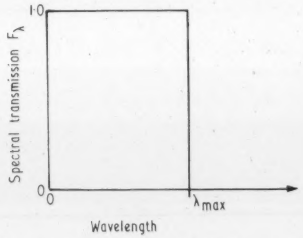


Fig. 3. Assumed transmission for sharp cut-off filter.

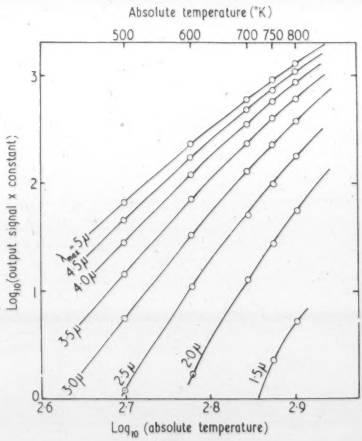


Fig. 4. Variation of output signal with temperature for sharp cut-off filter.

Table 1. *Approximate values of n for cut-off filters*

λ_{max} (μ)	Temperature (°K) 500	600	700	750
5·0	7·4	6·4	6·0	
4·5	7·6	7·0	6·2	
4·0	8·4	7·6	6·8	6·6
3·5	9·8	8·2	7·4	7·4
3·0	11·2	9·0	8·2	8·0
2·5		10·6	9·6	9·4
2·0		12·4	12·4	11·4

The value of n increases as the wavelength cut-off λ_{max} decreases. It also decreases with increasing temperature when the wavelength of peak emission from the source approaches or is less than λ_{max}. To obtain a pyrometer with a low emissivity error the overall spectral response should therefore be chosen so that the pyrometer is insensitive to the longer wavelengths, for example, wavelengths longer than 1·5 to 2·0 μ.

If the filter transmission F_λ (or spectral response R_λ) does not decrease from a maximum to zero at λ_{max}, but decreases steadily between wavelengths λ_1 and λ_2 as shown in Fig. 5, the corresponding output curves are as shown in Fig. 6. Table 2 gives the approximate values of n. Within a limited range of temperatures an extremely high value of n is obtained. For example, when F_λ (or R_λ) decreases steadily between $\lambda_1 = 2{\cdot}0\ \mu$ and $\lambda_2 = 3{\cdot}5\ \mu$, $n \simeq 17$ for temperatures between 700 and 750° K.

As more infra-red detectors and filters become available, it should therefore become possible to devise pyrometers with a low emissivity error over limited ranges of temperature.

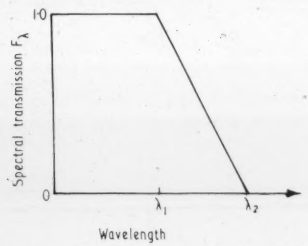


Fig. 5. Assumed transmission for sloped cut-off filter.

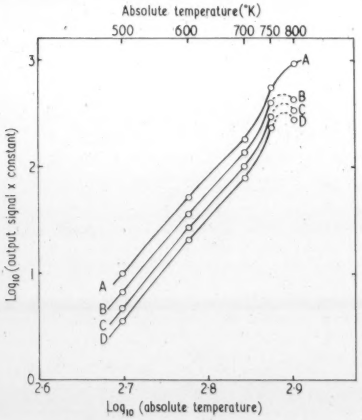


Fig. 6. Variation of output signal with temperature for sloped cut-off filter.

Curve A, $\lambda_1 = 3{\cdot}0\ \mu$, $\lambda_2 = 3{\cdot}5\ \mu$; curve B, $\lambda_1 = 2{\cdot}5\ \mu$, $\lambda_2 = 3{\cdot}5\ \mu$; curve C, $\lambda_1 = 2{\cdot}0\ \mu$, $\lambda_2 = 3{\cdot}5\ \mu$; curve C, $\lambda_1 = 1{\cdot}5\ \mu$, $\lambda_2 = 3{\cdot}5\ \mu$.

Table 2. *Approximate values of n with sloped cut-off filters*

Filter cut-off (μ) λ_1	λ_2	*Temperature* (° K) 500	600	700–750
3·0	3·5	8·7	8·7	~16
2·5	3·5	8·9	8·9	~16
2·0	3·5	9·0	9·0	~17
1·5	3·5	9·2	9·2	~15

Monochromatic radiation pyrometer

It is usual to express the relationship between the true and apparent temperatures measured with a monochromatic radiation pyrometer as

$$\left(\frac{1}{T}\right) - \left(\frac{1}{T_A}\right) = \lambda \ln \frac{\epsilon_\lambda}{c_2} \tag{8}$$

where λ is the operating wavelength, ϵ_λ the spectral emissivity at this wavelength and c_2 the radiation constant. For comparison with the preceding calculations it is interesting to express the output law for the monochromatic radiation pyrometer as a power law of absolute temperature with an exponent n which varies with temperature and wavelength. Using Planck's equation

$$E_\lambda d\lambda = c\lambda^{-5}\{\exp(c_2/\lambda T) - 1\}^{-1} d\lambda \tag{9}$$

and plotting $\log_{10} E_\lambda$ against $\log_{10} T$ for fixed wavelengths the curves shown in Fig. 7 are obtained. Again the slope of

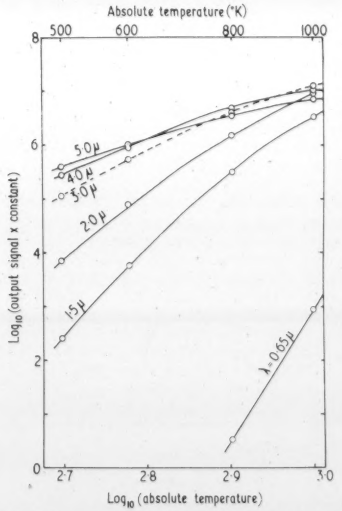


Fig. 7. Variation of output signal with temperature for monochromatic radiation pyrometer.

these curves gives the exponent n when the output law is expressed in the form

$$S = \text{const}\ T^n \tag{10}$$

and the fractional error in temperature measurement is

$$(T - T_A)/T = 1 - \epsilon^{1/n} \tag{11}$$

Table 3 gives approximate values of n, including a value for a disappearing filament pyrometer with an effective operating wavelength 0·65 μ at temperatures around 800° K, i.e. a temperature 200 deg K below the normal temperature range of this instrument.

It is again found that the exponent n increases as the operating wavelength of the pyrometer is decreased, and at a given wavelength decreases with increasing temperature. Comparing Tables 1 and 3 it appears that a pyrometer with uniform wavelength response to a wavelength λ_{max} leads to a slightly higher value of n than a monochromatic pyrometer operating at wavelength λ_{max}. It also has the practical advantage that it collects more radiation from the specimen.

Table 3. *Approximate values of n for a monochromatic radiation pyrometer*

Wavelength (μ)	*Temperature* (° K)			
	500	600	700	800
5	5·7	5·0	4·5	4·0
4	6·7	6·0	5·5	4·5
3	9·0	8·0	7·5	6·5
2·5	10·5	9·7	8·5	7·5
2·0	13·5	12·0	10·2	8·7
1·5	20·0	15·7	13·7	11·7
0·65				~24

3. Discussion

It is clearly shown that for temperatures in the range 500–800° K, a partial or monochromatic radiation pyrometer operating at wavelengths less than 5 μ gives a lower emissivity error than a total radiation pyrometer. The shorter the operating wavelengths of the pyrometer the smaller this error becomes.

In practice, the minimum operating wavelength that can be used is restricted by the detector noise and the minimum temperature to be measured. The detector noise determines the minimum measurable energy and the wavelength response the minimum temperature at which this energy can be collected. Since only a small fraction of the total energy emitted by surfaces between 500 and 800° K emerges at wavelengths less than 3·0 μ, it is necessary to compromise between obtaining a low emissivity error and devising a sensitive pyrometer.

It is interesting to note that if a pyrometer obeys a high power output law, the effect of ambient temperature upon the net energy exchange between the detector and source is reduced, although the change in the sensitivity of the detector with ambient temperature is unaltered.

Although in the temperature range 500–800° K a partial radiation pyrometer operating in the near infra-red region yields a slightly higher power law than a monochromatic radiation pyrometer operating at wavelength λ_{max}, for temperatures above 1000° K, when the specimen emits visible radiation, the disappearing filament pyrometer remains the most reliable radiation method of measuring temperature. Pyrometers with infra-red detectors and filters are therefore likely to be most useful for temperatures below 1000° K, where at present total radiation pyrometers are usually used.

A particularly difficult problem in temperature measurement arises when radiation methods must be used to measure the temperature of surfaces of low or variable emissivity. This problem arises, for example, when attempting to measure the temperature of moving aluminium during extrusion. A contact pyrometer cannot be used because it is liable to damage the surface of the soft aluminium and only a radiation pyrometer can overcome this difficulty. But since the emissivity of aluminium is low and variable (Reynolds 1961), the pyrometer must be chosen to have a low emissivity error.

The emissivity of a metal surface depends upon the surface finish and state of oxidation, and it is characteristic of low emissivity materials that variations in these factors cause proportionately large variations in emissivity. The spectral emissivities of most metals decrease with increasing wavelength (Price 1947), and it is therefore an added advantage to select a pyrometer which is only sensitive to wavelengths in the near infra-red. Low emissivity is also associated with high reflectivity, so that reflected radiation from nearby hot bodies may cause spurious signals, but this difficulty can usually be overcome by suitable shielding of the specimen.

When the emissivity is low $\epsilon^{1/n}$ approaches unity slowly as n increases (Fig. 8), and as ϵ falls it progressively becomes more difficult to reduce the emissivity error by increasing n. Nevertheless, it can be shown both theoretically and experimentally that by increasing n from 4 to say 11, there is a significant reduction in this error even for low emissivity materials.

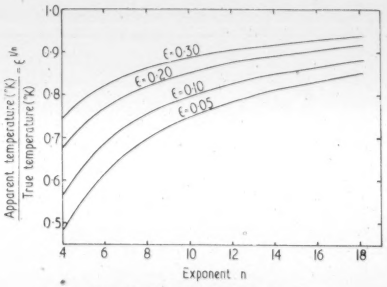


Fig. 8. Plot of $\epsilon^{1/n}$ against n.

A commercial lead sulphide cell pyrometer (Birnstingl 1958), in which the effective cut-off wavelength 2·5 μ was determined by a glass lens which focused radiation on to the photocell, was calibrated against a heated iron sheet of total emissivity 0·86. Under this condition of calibration the relationship between the true and apparent temperature becomes

$$0{\cdot}86T_A^n = \epsilon T^n \tag{12}$$

and the fractional error is given by

$$(T - T_A)/T = 1 - (\epsilon/0{\cdot}86)^{1/n}. \tag{13}$$

When this instrument was used to measure the temperature of aluminium specimens of total emissivity 0·20 and 0·30, the measured errors in temperature at 600° K were as shown in Table 4.

Table 4. *Temperature errors at* 600° K

Pyrometer	*Total emissivity*	*Measured error* (%)	*Calculated n*	*Calculated error* (%)
Lead sulphide	0·20	10	10·6	13
	0·30	6	10·6	10
Total radiation	0·20	19	4	31
	0·30	16	4	23

At 600° K when $\lambda_{max} = 2{\cdot}5\ \mu$, Table 1 gives $n = 10{\cdot}6$. The expected error can then be calculated from Eqn (13), and is also shown in Table 4. The difference between the measured and calculated errors is partly due to the increase in the spectral emissivity of aluminium with decreasing wavelength (Reynolds 1961), and partly arises because the pyrometer was sighted obliquely on the specimen at an angle of 80° to the normal. The emissivity of most surfaces tends to increase at low angles of emission.

For comparison this experiment was repeated using a 'total' radiation pyrometer, in which the focusing lens was stated to transmit radiation between $0{\cdot}3$ and $10{\cdot}0\ \mu$. In this case there was poor agreement between the measured and calculated errors, as may be seen from Table 4. The sighting angle was again 80° from the normal to the specimen surface. Nevertheless, these results show the advantage of using a pyrometer which exhibits a high power output law. If n could be increased further to say 17, the predicted error for these surfaces is reduced to 9% when $\epsilon = 0{\cdot}20$ and 7% when $\epsilon = 0{\cdot}30$.

Whilst considering 'total' radiation pyrometers it is worth noting that this term is at present loosely applied to pyrometers with mirror optics and those with lens optics. When lens optics are used with a total radiation receiver, the lens effectively defines the overall wavelength response. It is usual in pyrometers designed for use in the temperature range 500–800° K to use calcium fluoride lenses which transmit radiation to $10\ \mu$, but in instruments designed for a higher temperature range, silica or Pyrex glass lenses, which only transmit radiation to $3{\cdot}8$ and $2{\cdot}7\ \mu$ respectively, may be used.

With specimens of low and variable emissivity it is often better to calibrate the pyrometer against an actual specimen. When, for example, the lead sulphide cell pyrometer was calibrated for $\epsilon = 0{\cdot}20$ and used to measure the temperature of the specimen of total emissivity $0{\cdot}30$, the predicted and measured errors at 600° K were only 5% and $3\frac{1}{2}$% respectively. If n could be increased to 17 the predicted error becomes $2\frac{1}{4}$%.

Although the use of a detector with a high power output law is unlikely to completely solve the problem of using radiation pyrometers to measure the temperature of low emissivity materials in the range 500–800° K, it approaches a complete solution for materials with emissivities between $0{\cdot}5$ and $1{\cdot}0$. For lower emissivity materials the errors are still considerable, but it should be possible to reduce the emissivity error still further by combining a detector system which obeys a high power output law with a reflector method of increasing the effective emissivity. When a reflector is placed close to the specimen surface, it forms, with the specimen, an almost closed cavity. The multiple reflections which occur in this cavity increase the effective emissivity of the specimen. A pyrometer using this principle has already been developed by Land and Barber (1954) for temperature ranges extending from 50 to 1300° C. By calibrating on a surface of total emissivity $0{\cdot}8$, the emissivity error is less than 1% for total emissivities between $0{\cdot}6$ and $1{\cdot}0$.

Pattison (1959) has also developed instruments combining a total radiation detector with a hemispherical reflecting cavity, and a disappearing filament pyrometer with a hemispherical cavity, for measuring the temperatures of ingots in the range 800–1300° C. A combination of a reflecting cavity method with a detector showing a high power output law would be of value for low emissivity materials in the range 500–800° K.

References

AVERY, D. G., GOODWIN, D. W., and RENNIE, A. E., 1957, *J. Sci. Instrum.*, **34**, 394.

BIRNSTINGL, D. W., 1958, *Electronic Engng*, **30**, 187.

EMSLIE, A. G., and BLAU, H. H., 1959, *J. Electrochem. Soc.*, **106**, 877.

GILL, T. P., 1957, *J. Opt. Soc. Amer.*, **47**, 1000.

HARMER, J. D., and WATTS, B. N., 1955, *J. Sci. Instrum.*, **32**, 167.

LAND, T., and BARBER, R., 1954, *J. Soc. Glass Technol.*, **38**, 45.

LARSON, B. M., and SHENK, W. E., 1940, *J. Appl. Phys.*, **11**, 555.

PATTISON, J. R., 1959, *J. Iron Steel Inst.*, **191**, 163.

PRICE, D. J., 1947, *Proc. Phys. Soc.*, **59**, 118.

PYATT, E. C., 1954, *Brit. J. Appl. Phys.*, **5**, 264.

REYNOLDS, P. M., 1961, *Brit. J. Appl. Phys.*, **12**, 111.

Bandwidth of a moon communication circuit

by J. V. EVANS,*, B.Sc., Ph.D., University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Manchester

MS. received 6th March 1961

Abstract

Radar studies of the moon by a number of workers have shown that the principal reflections occur from a region at the centre of the visible disk having a radius of about one-tenth of that of the moon. Thus it is possible to use the moon as a reflector in a communication system, and several successful systems of this type have been demonstrated. This paper describes measurements made at Jodrell Bank to determine the single channel bandwidth of the system using double sideband amplitude modulated transmissions. It is found that the demodulation of the signals caused by the multiple scattering at the moon's surface restricts the bandwidth to about ± 1 *kc/s, although the demodulation does not appear to increase rapidly at higher modulation frequencies.*

1. Introduction

THE possibility that the moon might serve as a suitable reflector for a communication system was first examined in a theoretical paper by Grieg, Metzger and Waer (1948). These authors investigated the extent to which modulated signals would become demodulated as a consequence of multiple reflections from different surface elements of the moon. They argued that if the moon were to reflect radio waves in a similar manner to the way in which it reflects light, then modulation components above about 50 c/s in frequency would be barely detectable in the returned signals, in any circuit where the whole of the moon is illuminated by the aerial systems employed. This is because the brightness distribution of light across the moon's disk is approximately uniform (Markov 1948) and consequently signals would be returned from regions near the limbs of the moon. These would carry the modulation intelligence shifted in phase relative to the signals returned from the centre, giving rise to destructive interference. To achieve the necessary reduction of this effect using a narrow beam aerial system, a beamwidth of about 3 minutes of arc would be required, and this could be obtained only from an aperture some 1000 wavelengths in diameter. Even at the present time there are few steerable aerial systems which approach this size.

In recent years it has been shown (Evans 1957, Trexler 1958) that the brightness distribution over the moon's disk for reflected radio waves at metre and decimetre wavelengths is far from uniform. About 50% of the reflected energy is returned from a region at the centre of the visible disk having a radius of only $\frac{1}{10}$ of the total radius of the moon. Such a region has a depth which introduces a delay of only 60 μs. Signals returned from the limbs would be delayed 11·6 ms relative to those returned from the centre of the moon. The importance of this effect in improving the expected bandwidth of a moon communication system was immediately recognized and several groups have made experiments on the transmission and reception of modulated signals. This paper describes some experiments on the measurement of the bandwidth of a lunar communication circuit.

2. Theory for the bandwidth of a moon communications system

Because the moon is a target of considerable depth, signal components will be returned which have suffered different amounts of delay. The time interval between the first and last signals for an impulse reflected by the moon is 11·6 ms (corresponding to the delay of the pulse travelling from the leading edge to the limbs and back). The way in which the amplitude A_τ of the reflected pulse varies as a function of delay corresponds to the infinite impulse response of the system. That is, A_τ the spectrum of the delays observed when a short transmitter pulse is reflected from the moon can be thought of as the 'ringing' of an electrical filter when excited by a sudden impulse. At very low frequencies (less than 10 kc/s) the moon may be regarded as a two-port electrical filter whose bandpass characteristic is simply the transform of its infinite impulse response A_τ. Such frequencies, however, propagate through the earth's ionosphere only as 'whistling atmospherics' and cannot be used to study the moon. At considerably higher frequencies, where the wavelength has become much shorter than the depth of the effective scattering region, the average response of this electrical filter is largely independent of frequency. Consequently very little can be learned by studying the variation of echo amplitude as a function only of a single frequency. Grieg, Metzger and Waer (1948) assumed that by using an amplitude modulated signal, the multiple scattering from the various scattering centres would cause the signal to appear to be demodulated on reflection, in an identical manner to the frequency response characteristic of the moon at very low frequencies, that is, the function $A(\omega_m)$ which relates the intensity of the modulation components after detection in the receiver to the modulation frequency ω_m, would also be the transform of the infinite impulse response A_τ. This assumption does not seem to be justified as it is based upon an argument which assumes that the instantaneous phase of the carrier wave can be neglected. Alternatively if an amplitude modulated wave be thought of in terms of its spectrum (a single tone of frequency ω_m gives rise to sidebands spaced $\pm\omega_m$ about the carrier frequency), it is no longer apparent why the intensity of the beat product of these three reflected frequencies should be simply given by the transform of the infinite impulse response. In view of the absence of any satisfactory alternative theory which would relate the

* Now at Lincoln Laboratory, M.I.T., Lexington, 73, Massachusetts, U.S.A.

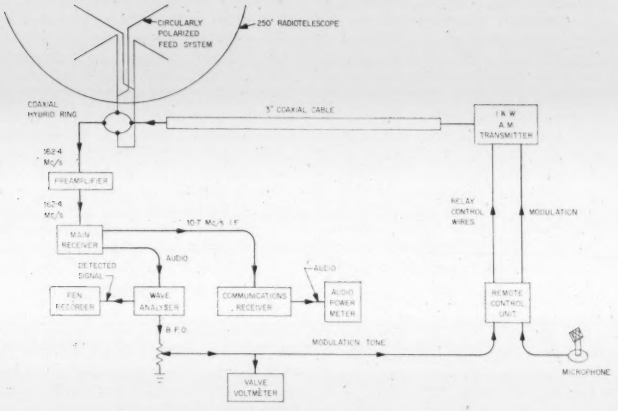


Fig. 1. Arrangement of equipment used at Jodrell Bank to determine the demodulation of an a.m. signal caused by reflection from the moon.

demodulation function $A(\omega_m)$ to the mean distribution of echo amplitude with delay A_τ, an experiment was planned at the Nuffield Radio Astronomy Laboratories to determine this function for amplitude modulated signals.

3. The observations

The experimental work reported here was conducted in April 1960. In this work amplitude modulated signals were transmitted towards the moon, and the characteristics of the echoes were then examined. Earlier, in July 1959, a similar experiment was undertaken using a narrow band frequency modulation system, but provided results which were inconclusive (Evans 1959), partly because a suitable wave analyser was not then available. In the later work the 250 ft steerable telescope was used together with a 1 kw high level modulation a.m. transmitter lent by Messrs. Pye Ltd., of Cambridge, England. The parameters of the equipment were: frequency 162·4 Mc/s, transmitter power 900 w (transmitter unmodulated), aerial gain 40 dB, polarization right circular, noise factor 8, main receiver i.f. bandwidth 8 kc/s, and feeder losses 2 dB.

The arrangement of this equipment is shown in Fig. 1. The mean signal-to-noise ratio observed at the output of the communications receiver shown in this figure was 10·3 ± 2 dB, for a tone at a frequency of 1 kc/s applied at a level sufficient to modulate the transmitter 100%. This result was obtained by recording the readings of a thermocouple power meter connected across the loudspeaker terminals of this receiver for many 2·5-second transmissions. This receiver had a bandwidth of 3 kc/s, and consequently the noise power referred to the input of the pre-amplifier is −160 dBw. The expected signal-to-noise ratio may be calculated as: P_s/P_n = path loss + feeder loss + aerial gain + aerial aperture + transmitter power − receiver noise power. The path loss is here defined as the attenuation introduced by the circuit when measured between an isotropic transmitting aerial and a receiving aerial of unit aperture. When defined in this way the path loss is not a function of the wavelength, and the mean of the independent observations of many different workers is −247 dB (Evans 1960).

Thus the expected signal-to-noise ratio is

$$P_s/P_n = (-247 - 2 + 40 + 35 + 30 + 160)\text{dB}$$
$$= +16 \text{ dB}.$$

This is the expected signal-to-noise ratio of the carrier at the input to the receiver. For a carrier wave which is 100% amplitude modulated, the power in the sidebands is 3 dB less than that of the carrier, and consequently we should expect that at the output of the receiver the signal-to-noise ratio of the modulation tone would be +13 dB. The observed signal was approximately 3 ± 2 dB less than this.

The audio tone used to modulate the transmitter was derived from a beat-frequency oscillator built into a wave analyser. The wave analyser was connected to the audio output of the main receiver and this arrangement made it automatic that the wave analyser would be tuned to the tone used to modulate the transmitter. The half power bandwidth of the wave analyser was 2·5 c/s and consequently the mean signal-to-noise ratio of a 1 kc/s tone at the output of the receiver was increased to about +34 dB. The wave analyser contained a linear detector, and the voltage developed by this was used to operate a high speed pen recorder. The transmitter could not be modulated successfully at frequencies substantially lower than 300 c/s or above 10 kc/s, due to the design of the modulator equipment. As a result measurements were made only at frequencies of 300 c/s, 500 c/s, 750 c/s, 1 kc/s, 2 kc/s, 3 kc/s, 4 kc/s, 6 kc/s, 8 kc/s and 10 kc/s.

The procedure adopted was to pulse the transmitter on and off at 3-second intervals, whilst the radio telescope was continuously following the moon. Some 240 echoes were obtained at each of the modulation frequencies listed above. The amplitude of the echo was recorded on paper chart, and subsequently read off at three equally spaced intervals along each echo. In this way some 720 values for echo amplitude were obtained for each of the ten modulation

frequencies. The mean and the r.m.s. deviation were then calculated, and finally corrections were applied for the bandpass characteristics of the receiver. It is necessary to average over a large number of echoes, because deep fading of the echoes is observed as a consequence of the librational spin of the moon. This is the apparent rotation of the moon caused chiefly by the diurnal motion of the observer on the earth. The other causes of libration and the rates of rotation which they introduce have been discussed by Browne *et al.* (1956). The results of the experiments are shown in Fig. 2,

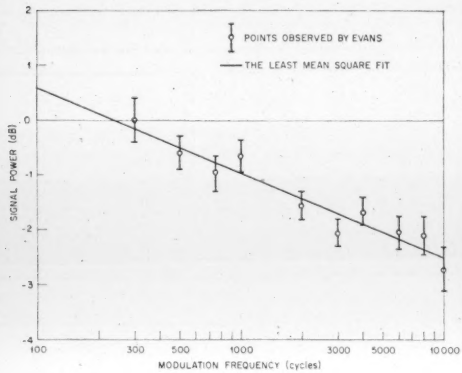


Fig. 2. Variation of modulation signal intensity as a function of frequency for an a.m. signal reflected by the moon.

together with a straight line fitted by 'least mean squares' procedure.

Great care was taken to adjust the i.f. and audio circuits of the receiver so that the response was substantially independent of frequency over the range 100 c/s–7 kc/s. For frequencies above 7 kc/s the response was adjusted such that it fell rapidly being about 7 dB down at 10 kc/s. This was necessary to keep the noise voltage appearing at the input terminals of the wave analyser from overloading this instrument. For each measurement the modulation level at the transmitter was adjusted to be 75%. The measurements for any one modulation frequency were made over several periods of 5 minutes separated at hourly intervals so that any Faraday rotation (caused by a residual amount of plane polarized waves which might be radiated by the aerial feed system) would tend to be averaged out. With these precautions it is believed that the systematic errors in the measurements have been reduced to the same order as the uncertainty caused by the variation of the echo amplitude, i.e. less than $\pm 0{\cdot}4$ dB.

The points in Fig. 2 have been normalized relative to the value observed at 300 c/s modulation frequency. However, the slope of the straight line which has been fitted to the results suggests that there is already some demodulation at 300 c/s, and it was unfortunate that measurements could not be made at a lower frequency. It is unlikely that a linear relation exists between the demodulation and the modulation frequency below 300 c/s, and hence it would not be justifiable to extrapolate the observed results to find how much demodulation exists at this frequency. The observed signal-to-noise ratio at 1 kc/s was 3 ± 2 dB lower than expected. The error in this measurement represents both the scatter of the observed intensities and the uncertainty in some of the equipment parameters. Thus it is concluded that the half power bandwidth of the system is approximately ± 1 kc/s. The useful bandwidth is considerably larger than this because there is only a further 2 dB fall in signal intensity over the range of modulation frequencies 1 kc/s to 10 kc/s.

No attempt was made to measure the harmonic distortion of the signal at the output of the receiver. The reasons for this were (*a*) the harmonic content of the reflected signal is difficult to measure for frequencies above about 3 kc/s, because the frequencies which give rise to this distortion move outside the passband of the receiver; (*b*) since the signal-to-noise ratio of the fundamental in the wave analyser is only about +30 dB measurements of the harmonic content would be subject to large uncertainties due to the presence of noise.

4. Discussion

In this paper the bandwidth of the moon has been defined by analogy with an electrical circuit as being the passband of the system. An alternative definition for a time-varying network such as this would be the frequency separation between two carrier waves which causes their amplitudes to become uncorrelated. This function has been measured by Ingalls *et al.* (1961), who find that for a frequency separation $\Delta\omega$ of 1 kc/s between the carrier waves the correlation of their amplitudes $\gamma(\Delta\omega)$ is reduced to 0·25. Thus it is clear that these two definitions yield approximately the same results. Hagfors (1960) has examined the relationship between the correlation $R(\Delta\omega)$ which exists between the complete waveforms of two carrier waves and the average impulse response of the circuit A_τ. He showed that provided that there are many scattering elements, so that the returned signals resemble narrow-band gaussian noise, A_τ and $R(\Delta\omega$ are related by

$$\overline{A_\tau^2} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \exp(i\Delta\omega\tau) R(\Delta\omega) d(\Delta\omega)$$

and

$$R(\Delta\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \overline{A_\tau^2} \exp(i\Delta\omega\tau) dt.$$

Thus the correlation coefficient $R(\Delta\omega)$ for two frequencies separated by $\Delta\omega$ is given by the Fourier transform of the average infinite impulse power response $\overline{A_\tau^2}$. In most experiments, however, the complete r.f. waveforms are not examined but only their amplitudes. Ingalls *et al.* (1961) determined the correlation between the detected signals, and employed linear detectors at the outputs of their two receivers. If $|G(\omega_1)|$ is the amplitude of the envelope of the signal at a frequency ω_1, then we require the average of the product $|G(\omega_1)|\,|G(\omega_2)|$. Price (private communication) has suggested that this will not differ appreciably from the average observed using square law detectors, i.e.

$$\overline{|G(\omega_1)|\,|G(\omega_2)|} \simeq \overline{|G(\omega_1)|^2\,|G(\omega_2)|^2}$$

and that this is given by

$$\overline{|G(\omega_1)|^2\,|G(\omega_2)|^2} = \int_{-\infty}^{+\infty} \overline{A_\tau^2} dt + \left[\int_{-\infty}^{+\infty} \overline{A_\tau^2} \exp\{i\tau(\omega_1 - \omega_2)\} dt\right]^2$$

assuming that ω_1 and $\omega_2 \gg 1/\text{duration of } \overline{A_\tau^2}$ and that the signals resemble narrow-band Gaussian noise. The first term on the right of the above equation is simply a 'd.c. term' representing the small value for the correlation that will exist at widely spaced frequencies. Thus Price (private communication) concludes that the normalized correlation function $\gamma(\Delta\omega)$ observed between the amplitudes of the two signals at frequencies ω and $\omega + \Delta\omega$, is given by the square of the Fourier cosine transform of the average infinite impulse power response $\overline{A_\tau^2}$.

Because the moon is a solid target spinning with a uniform angular velocity it would also be expected that a relation will exist between the function for the demodulation observed here (Fig. 2) and the power response $\overline{A_\tau^2}$. The relationship for a system where one sideband is suppressed is straightforward. For such a system the amplitude of the modulation tone ω_m corresponding to a frequency separation between the carrier and sideband $\omega_m = \omega_1 - \omega_2$ is simply $|\overline{G(\omega_1)}|\,|\overline{G(\omega_2)}|$. That is the demodulation function has the same form as the autocorrelation function $\gamma(\Delta\omega)$, but whereas it is customary to remove the average in $\gamma(\Delta\omega)$, i.e.

$$\gamma(\Delta\omega) = \frac{|\overline{G(\omega)}|\,|\overline{G(\omega + \Delta\omega)}| - |\overline{G(\omega)}|^2}{|\overline{G^2(\omega)}| - |\overline{G(\omega)}|^2}$$

(so that $\gamma(\Delta\omega) \to 0$ as $\Delta\omega$ becomes large), the demodulation function $A(\omega_m)_{\text{sup. sideband}}$ is simply

$$A(\omega_m)_{\text{sup. sideband}} = |\overline{G(\omega)}|\,|\overline{G(\omega + \omega_m)}|.$$

Where double sideband a.m. signals are transmitted, the phases as well as the amplitudes of the two reflected sidebands are important. This is because it is possible for the two sidebands to beat with the carrier in the receiver and generate tones which have the same frequency but not the same phase. Thus the demodulation function $A(\omega_m)_{\text{a.m.}}$ becomes the average of the vector sum of two beat products $G(\omega)G(\omega + \omega_m)$ and $G(\omega)G(\omega - \omega_m)$

$$A(\omega_m)_{\text{a.m.}} = \overline{G(\omega)G(\omega + \omega_m) + G(\omega)G(\omega - \omega_m)}.$$

It is not clear what relation $A(\omega_m)_{\text{a.m.}}$ has to $A(\omega_m)_{\text{sup.sideband}}$. When considered qualitatively it can be seen that two opposing effects will take place for high values of ω_m:

(*a*) The sidebands $G(\omega + \omega_m)$ and $G(\omega - \omega_m)$, being separated by $2\omega_m$, soon cease to be well correlated. As a result when either one is weak there is no strong probability that the other will be weak also, and because only one sideband is required to generate the modulation tone, the likelihood of the resultant audio tone being momentarily poor is reduced.

(*b*) An opposing effect is the fact that when both sidebands are strong they will generate tones which are not in phase. The resulting signal will be weaker than that generated by either one separately when they differ in phase by more than 90°. This will reduce the average signal level and cause a double sideband a.m. system to have a smaller useful bandwidth than a suppressed sideband system.

5. Summary

It has been shown that the alternative definition of the bandwidth of the circuit, by means of the correlation–frequency separation function, is directly related to the impulse response of the network. However, no similar simple relationship has been found for the demodulation function $A(\omega_m)$ measured here (Fig. 2). Thus it is not possible to compare this measurement with the results of other radio-echo studies of the moon.

A double sideband a.m. system will exhibit a deterioration in the level of the modulation as the modulation frequency is increased. This effect causes the channel to have a half power bandwidth of about ±1 kc/s, but because the reduction in the level of the modulation does not appear to increase rapidly with frequency a considerably larger bandwidth could, in theory, be used. For instance, a channel employing modulation frequencies of 300 c/s to 10 kc/s would exhibit a response at the high frequencies only 3 dB worse than at the low frequencies. A compensating network could be employed either at the transmitter or at the receiver to remove this effect. However, as Ingalls *et al.* (1961) have observed the useful channel bandwidth is restricted to the range of modulation frequencies, which generate sidebands that are well correlated, and this is only of the order of 1–2 kc/s. If a larger channel width is employed it becomes impossible to use automatic gain control to remove the fluctuations in signal level, and the intelligence sounds distorted because the relative levels of the individual tones in the reflected signal will be different from the levels in the transmitted waveform.

6. Conclusion

The moon can be used as a reflector in a communication system which has adequate bandwidth for transmitting speech intelligence, but a single channel would be inadequate for music. A system of frequency diversity would seem to be the best method to overcome the rapid fading caused by the moon's libration. Evans (1957) proposed that a system using narrow-band f.m. could achieve this effect, but single sideband or conventional a.m. systems make better use of the available bandwidth.

Acknowledgments

I should like to thank Professor Sir Bernard Lovell, the Director of the Nuffield Radio Astronomy Laboratories, for the use of the facilities there. I am indebted to my colleagues Messrs. G. N. Taylor and A. C. Prior for much assistance with the experiment. I also wish to thank Dr. R. Price of Lincoln Laboratory, M.I.T., who has greatly assisted my understanding of the characteristics of time-varying networks. Finally, I should like to thank Messrs. Pye, Ltd., for the loan of the transmitter. Some of the research reported in this paper was supported by the Electronics Research Directorate of the Air Force Cambridge Research Laboratories under Contract No. AF 61(052)-172.

References

Browne, I. C., Evans, J. V., Hargreaves, J. K., and Murray, W. A. S., 1956, *Proc. Phys. Soc. B*, **69**, 901.

Evans, J. V., 1957, *Proc. Phys. Soc. B*, **70**, 1105.

—— 1959, *Technical Status Report TSR* 3, *of Contract No.* AF 61(052)-172.

—— 1960, *Lincoln Laboratory Report* 3*G*-0003.

Grieg, D. D., Metzger, S., and Waer, R., 1948, *Proc. Inst. Radio Engrs, N.Y.*, **36**, 652.

Hagfors, T., 1960, *Scientific Report No.* 8 *of Contract* AF 19(604)-2193.

Ingalls, R. P., Bird, L. E., and Day, J. W. B., 1961, *Proc. Inst. Radio Engrs, N.Y.*, **49**, 631.

Markov, A. V., 1948, *Astr. J. Moscow*, **25**, 172.

Trexler, J. H., 1958, *Proc. Inst. Radio Engrs, N.Y.*, **46**, 286.

Production of accurate gas and vapour mixtures

by D. W. HILL, M.Sc., A.M.I.E.E., A.Inst.P., Research Department of Anaesthetics, Royal College of Surgeons of England, London, W.C.2

MS. received 10*th March* 1961, *in revised form* 4*th April* 1961

Abstract

With the increasing use of physical methods of gas and vapour analysis, there is a need for a convenient means of calibration. It is confirmed that the use of gas and vapour mixtures stored in cylinders under pressure is a suitable method. Such mixtures are stable with time. The analysis of gas and vapour mixtures has been found to agree to within 1% *of the calculated value.*

Introduction

In laboratories concerned with gas and vapour analysis, increasing use is being made of physical methods such as mass spectroscopy, infra-red gas analysis and gas chromatography. For quantitative work such analysers must be calibrated against accurately prepared mixtures of known composition. Such a situation arises in modern anaesthetic research where increasing use is being made of the technique of gas chromatography (Adlard and Hill 1960, Hill 1960). The great advantage of the gas chromatograph lies in the fact that on the one instrument may be analysed all the gases and vapours encountered in anaesthetic practice. The components to be analysed include oxygen, nitrogen, carbon dioxide, nitrous oxide, cyclopropane, ether, halothane (2-bromo-2-chloro-1, 1, 1-trifluorethane), chloroform and trichloroethylene.

Several methods are available for the production of known gas and vapour mixtures. Gas mixtures may be prepared by mixing together streams of gases, the individual volume flows being measured with 'rotameter' type flow meters. However, unless complicated regulating equipment is used it becomes somewhat inconvenient to hold the flows steady. Known vapour concentrations may be prepared by vaporizing a known mass of the liquid per unit interval of time into a known volume flow rate of the gas (Jahn 1955).

Alternatively, the liquid to be vaporized can be held in a metal saturating vessel, which is cooled to a known temperature in a solid carbon dioxide–petroleum ether bath. A steady flow of some 50 ml/min of the required carrier gas is drawn through the vessel. A knowledge of the barometric pressure and the variation of the saturation vapour pressure of the liquid with temperature enables the concentration of the vapour emerging from the vessel to be calculated (Edmondson 1957). It is essential to prevent the occurrence supersaturation of the issuing vapour. This is accomplished by passing the vapour stream through a metal spiral immersed in the coolant bath. Rubber connections should not be used since they can readily absorb vapours. Care must also be taken to ensure that gas and vapour streams are thoroughly mixed.

The most generally useful method of calibrating gas and vapour analysers is to employ accurately prepared mixtures contained in metal cylinders under pressure. Reproducible calibrations are readily obtainable and the cylinders may be used in portable equipment away from laboratory facilities.

Cylinder filling procedure

The cylinders to be used should be fitted with a competent fine control valve and must be able to withstand safely the required final filling pressure. If made of steel they must be free from rust on the inside surfaces. Rust may absorb anaesthetic vapours or cause a diminution in the oxygen concentration (Haldane 1920). To avoid this difficulty it is best to have the inside surfaces of the cylinders cleaned by shot-blasting and then electro-tinned. The cylinder filling manifold is constructed in two parts, of which the low-pressure part is shown in Fig. 1. By means of a rotary oil pump the cylinder and manifold are evacuated to a pressure of 0·1 mm Hg or better. For cylinders to be filled with mixtures containing components in the parts per million range, an oil diffusion pump is used with the manifold pressure measured on a Penning gauge. Since a cylinder to be evacuated may have previously contained a vapour mixture, the removal of this vapour is followed with a Pirani gauge. Once the manifold and cylinder have been evacuated the vacuum needle valve is opened and shut several times to pump off any air present above the liquid to be vaporized. The pump is then shut off from the system and vapour admitted by gently warming the liquid with a hot-air blower. The actual vapour pressure in the system is read off on the mercury manometer, due allowance being made for level changes occurring in the reservoir. The actual pressure of vapour admitted is chosen to be below the saturation vapour pressure of the liquid at the lowest ambient temperature to which the cylinder is likely to be subjected in use. In this way there is no fear of the vapour condensing in the cylinder. After being filled, the cylinder is removed from the manifold and the vacuum pump oil purged with an air ballast flow to remove any vapour from the oil.

The cylinder is now attached to the high-pressure manifold shown in Fig. 2. The commercial gas cylinders containing the various components of the diluent gas mixture are connected to the manifold via a master fine-control valve. This ensures a fine control over the filling pressure irrespective of the condition of the individual cylinder valves. Low partial pressures are measured with a mercury column in parallel with a 0–20 lb/in^2 Bourdon tube gauge. Higher partial pressures are measured with a 0–500 lb/in^2 Bourdon tube gauge in parallel with a dead-weight tester. The oil filling of the dead-weight tester is prevented from entering the manifold by means of a neoprene diaphragm. The manifold is conveniently constructed from nylon pressure tubing.

Before a cylinder is filled, all unions are loosened and the

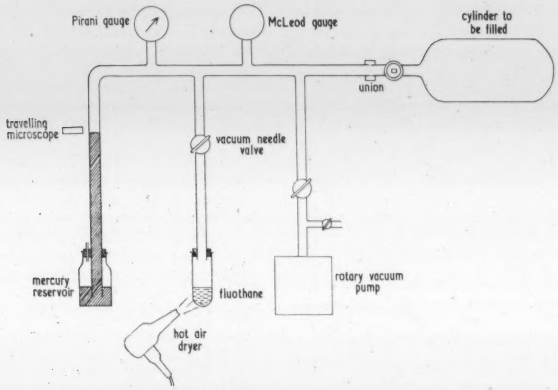


Fig. 1. Vacuum manifold for cylinder filling.

system thoroughly flushed with gas. With the unions tight, and the manifold pressure above atmospheric, the valve of the cylinder to be filled is slowly opened. The cylinder valve and the master valve on the manifold are then alternately opened in small steps until the cylinder has its valve wide open and is filled to the required pressure. The Bourdon tube gauges act merely as indicators, the pressures being measured accurately on either the mercury column or the

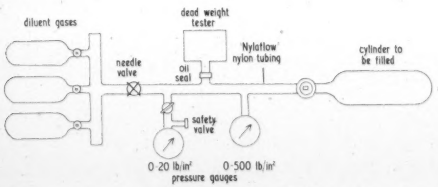


Fig. 2. Pressure filling manifold for cylinders.

dead-weight tester. When filled the cylinder is shut off from the manifold by closing the cylinder tap for at least one hour, to allow it to attain temperature equilibrium with its surroundings. At the end of the hour, when the tap is opened, the manifold and cylinder pressure will have fallen slightly and is then reset to its final value. The cylinder is then shut off and the manifold pressure released. If more than one gas is to be used in the filling the manifold is flushed with the next gas and the procedure repeated. Immediately the cylinder is filled it is labelled to record its contents. When cylinder mixtures are to be inhaled by human subjects it is most desirable that the composition of the mixture be checked by an analytical procedure.

In computing the cylinder contents it must be remembered that the dead-weight tester measures pressure with respect to atmosphere and not to a vacuum. Atmospheric pressure is read on a Fortin barometer in the same laboratory as the filling manifolds, and the height of the barometer converted to the equivalent height at 0° C. Use is then made of the relationship that a pressure of 14·70 lb/in^2 absolute is equivalent to the pressure due to a 760 mm column of mercury at 0° C. This method of preparing mixtures in cylinders can be applied to most gases and vapours. It is recommended that mixtures of inflammable substances such as ether be prepared in nitrogen rather than oxygen. By this means any risk of combustion is reduced.

Stability of cylinder mixtures

As previously described, the cylinder is left for at least one hour on the manifold before the pressure is finally adjusted. In the case of a mixture of carbon dioxide in oxygen it was found that the composition of the mixture, as determined on removing the cylinder from the manifold, subsequently remained constant. For ether and halothane mixtures in nitrogen the composition of the mixture became constant only after the lapse of some four hours after filling. It is the practice in this Laboratory to leave any cylinder overnight before using it. Thus the cylinders are not rolled to hasten the process of mixing, but we rely on thermal agitation to accomplish this. As an example of the effect of rolling, Bann (private communication) states that a cylinder filled to 1800 lb/in^2 with a mixture of 50% nitrogen and 50% oxygen by volume attained a constant composition after 135 minutes without rolling and in a horizontal position, after 8 minutes with rolling. It would seem that if a high-pressure cylinder is left to mix in a vertical position the process may take several days. Here the area across which diffusion can occur is much smaller than in the horizontal position.

It has been suggested (Astrup 1957) that mixtures of carbon dioxide in oxygen compressed in cylinders change their composition with time. The explanation for this is stated to be that the more dense carbon dioxide settles out at the bottom of the cylinder. From kinetic theory one would not expect this to happen, as under experimental conditions the effect of gravity on the gas mixture can be neglected. Even in the critical region the effect of gravity may not be significant. This has been demonstrated for ethylene by Maas and Geddes (1937). Two cylinders containing mixtures of carbon dioxide in oxygen were secured, one in a vertical position and the other in a horizontal position. Analyses of the carbon dioxide concentration were

Table 1. *Cylinder concentration stability (Nov. 1960–Jan. 1961)*

Cylinder left permanently horizontal (pressure 1900 lb/in² gauge) Nominal 5% CO_2 in O_2 (British Oxygen Co. Ltd.)

Date	23 Nov.	27 Nov.	5 Dec.	7 Dec.
CO_2 concentration (% v/v)	4·59	4·59	4·61	4·59
Date	12 Dec.	4 Jan.	12 Jan.	25 Jan.
CO_2 concentration (% v/v)	4·61	4·63	4·60	4·61

Cylinder left permanently vertical

Nominal 7% CO_2 in O_2 (British Oxygen Co. Ltd.)

Date	8 Dec.	13 Dec.	12 Jan.	30 Jan.
CO_2 concentration (% v/v)	6·73	6·71	6·74	6·73

Table 2. *Cylinder concentration stability (Nov. 1960–Jan. 1961)*

Cylinder deliberately run down

Nominal 5% CO_2 in O_2 (British Oxygen Co. Ltd.)

Date	14 Nov.	14 Nov.	17 Nov.	17 Nov.	21 Nov.
Pressure (lb/in² g)	1800	1440	1080	720	360
CO_2 concentration	5·21	5·20	5·19	5·21	5·15
Date	23 Nov.	28 Nov.	29 Nov.	7 Dec.	4 Jan.
Pressure (lb/in² g)	260	180	100	50	30
CO_2 concentration	5·19	5·19	5·17	5·19	5·19

made with a Lloyd–Haldane analyser (Lloyd 1958) at intervals of several days. The constancy of the results is shown in Table 1. Another cylinder was deliberately run down, over a period of several days, and the carbon dioxide concentration measured for various cylinder pressures. Table 2 shows that the maximum variation was 1·15% of the original value. Timms *et al.* (1958) report a significant change in the concentrations of impurities in carbon dioxide depending on the rate of withdrawal of the gas from the cylinder. They attribute the effect to a different distribution of impurity gases between the gas and the liquid phases in the cylinder. Vapour mixtures in cylinders have been found to be stable to within 1% of the original value using infra-red analysers of the Luft type (Luft 1943). The analysers were calibrated against an independent method based on saturation vapour pressures as previously described.

Compressibility corrections in cylinder filling

If the composition of the gas mixture in a cylinder is to be calculated accurately on a pressure basis, then it is necessary to be able to take into account the departures from the perfect gas laws of the gases used. Calculations based only on Dalton's law of partial pressures* (Dalton 1802) and Boyle's law may lead to error.

* It is interesting to note that Dalton postulated his law as a solution to the problem of why the atmosphere does not settle out into layers under the action of gravity. Modern work (Chackett *et al.* 1949) shows that there is no detectable difference between the composition of ground level and stratosphere air.

Gaseous solutions at moderate pressures behave as ideal solutions (Lewis and Randall 1923), and obey the law of additive partial volumes (Din, British Oxygen Co. Research Report No. E 2841 1960). For gases this law states that in the absence of chemical reaction, when gases at the same temperature and pressure are mixed, there is no change in the total volume nor in the pressure and temperature after mixing. This law does not apply to cylinders filled to such a pressure that any one of the components, if subjected alone to this final filling pressure, would liquefy. Under these conditions a form of compressibility correction based on Dalton's law is employed. The mole fraction of the desired component is calculated using experimentally determined values for the molar volumes.

Detailed evidence of the pressure and composition ranges over which a gas mixture will obey the law of additive partial volumes (due to Amagat) is not available.

The calculations set out in Table 3 show that a mixture computed on a simple Dalton's law and Boyle's law basis to

Table 3. *Compressibility correction for a 5% CO_2 in O_2 cylinder*

CO_2 filling pressure	= 20 lb/in² abs. = 1·36 atm
Final filling pressure	= 400 lb/in² abs. = 27·2 atm
Room temperature	= 20° C
Specific volume CO_2 at 20° C, 1·36 atm	= 17 580 cm³/mole
Specific volume CO_2 at 20° C, 27·2 atm	= 739 cm³/mole
Specific volume O_2 at 20° C, 27·2 atm	= 870 cm³/mole

Let V = total cylinder volume (in cm³)
V_1 = partial volume CO_2 in final mixture
V_2 = partial volume O_2 in final mixture

Number of moles CO_2 at 1·36 atm = $V/17\,580$
Number of moles CO_2 at 27·2 atm = $V_1/739$
Hence $V/17\,580 = V_1/739$
Number of moles O_2 at 27·2 atm = $V_2/870$
$V = V_1 + V_2$ (law of additive partial volumes)

$$\text{Mole fraction of } CO_2 = \frac{V/17\,580}{V/17\,580 + V_2/870}$$

$$= 0{\cdot}0491 = 4{\cdot}91\% \text{ v/v } CO_2$$

contain 5·00% v/v of carbon dioxide in oxygen at a filling pressure of 400 lb/in² absolute would actually contain 4·91%. It can be seen that the effect of the compressibility factor is to cause the cylinder to contain a greater mass of oxygen than was originally expected. The pressure, volume and temperature data were obtained from Din (1956) and Michels *et al.* (1954). Table 4 shows that the values for various mixtures calculated from chemical analysis by the method of Scholander (1947) agree to within 1%. The method can be extended to ternary mixtures as shown in Table 5. When a major component comprises a relatively compressible gas such as nitrous oxide, it is seen that the compressibility correction is by no means negligible. The analysis of mixtures containing nitrous oxide was performed with a Haldane analyser using saturated sodium hydroxide (Nunn 1958). The pressure, volume and temperature data for nitrous oxide were obtained from the University of Texas (Hirth, private communication). The data of Britton (1929) are not accurate.

Table 4. *Analysis of mixtures of CO_2 in O_2*

(1)	(2)	(3)	(4)
A	2·90 ± 0·03	2·93 ± 0·03	501
B	4·88 ± 0·03	4·89 ± 0·03	407
C	4·91 ± 0·03	4·95 ± 0·03	406
D	4·91 ± 0·03	4·94 ± 0·03	406
E	6·90 ± 0·04	6·91 ± 0·03	291
F	6·88 ± 0·04	6·83 ± 0·03	292

(1) Cylinder; (2) calculated % CO_2 v/v; (3) % CO_2 Scholander analyser; (4) cylinder pressure (lb/in^2 abs).

On a straight partial pressure basis, without any compressibility corrections, cylinder A would give 3·0% CO_2, cylinders B, C, D 5·0% CO_2 and cylinders E, F 7% CO_2.

Table 5. *Compressibility corrections for a 5% v/v CO_2 concentration in a diluent comprising* (70% N_2O, 30% O_2)

CO_2 filling pressure = 20 lb/in^2 abs. = 1·36 atm
O_2 filling pressure = 114 + 20 = 134 lb/in^2 abs. = 9 atm
N_2O filling pressure = 266 + 134 = 400 lb/in^2 abs.
= 27·2 atm

Let

A = spec. vol. cm^3/mole CO_2 at 20° c, 1·36 atm = 17 580
B = spec. vol. cm^3/mole CO_2 at 20° c, 9 atm = 2 539
C = spec. vol. cm^3/mole O_2 at 20° c, 9 atm = 2 649
D = spec. vol. cm^3/mole CO_2 at 20° c, 27·2 atm = 739
E = spec. vol. cm^3/mole O_2 at 20° c, 27·2 atm = 870
F = spec. vol. cm^3/mole N_2O at 20° c, 27·2 atm = 732

$$\%\ \text{v/v}\ CO_2 = 100\left\{1 + \frac{(A-B)}{C} + \frac{1}{F}\left[A - D - \frac{E(A-B)}{C}\right]\right\}^{-1}$$

$$= 4{\cdot}36\%$$

This was 4·40% by Haldane analysis.

From Dalton's and Boyle's laws, one could expect to produce a binary mixture of 100 parts per million by using 3·04 mm Hg pressure of the minor component in a total cylinder pressure of 40 atmospheres. For a 100 parts per million mixture of carbon dioxide in oxygen the compressibility calculation gives a result of 102 parts per million. For 100 parts per million of oygen in carbon dioxide the calculated result becomes 71 parts per million.

It is suggested that cylinders whose calculated compositions agree with the chemical analysis to within 1% be held as standards in the laboratory. These cylinders can then be used to check the constancy of the chemical analysers, since unexpected trouble may arise with these because of leaking taps and faulty reagents. In addition, the cylinders provide a useful cross-check on the absolute calibration of the chemical analyser.

For the past four years a variety of physical methods of gas analysis, including various infra-red analysers and gas chromatographs, have been regularly standardized against calibration cylinders. No unexpected results have been found.

Acknowledgments

Grateful acknowledgments must be paid to the advice on cylinder filling methods offered by the British Oxygen Co. Ltd., and the Infra-Red Development Co. Ltd. Dr. Din has contributed many helpful discussions on compressibility corrections. Miss C. Casselle and Dr. Nunn of this Department have kindly performed a large number of the chemical gas analyses.

References

ADLARD, E. R., and HILL, D. W., 1960, *Nature, Lond.*, **186,** 1045.
ASTRUP, P., 1957, *Scan. J. Clin. Lab. Invest.*, **33,** 8.
BRITTON, G. T., 1929, *Trans Faraday Soc.*, **25,** 522.
CHACKETT, K. F., PANETH, F. A., and WILSON, E. J., 1949, *Nature, Lond.*, **164,** 128.
DALTON, J., 1802, *Mem. Manchester Lit. Phil. Soc.*, **5,** 535.
DIN, F., 1956, *Thermodynamic Functions of Gases* (London: Butterworth).
EDMONDSON, W., 1957, *Brit. J. Anaesth.*, **24,** 28.
HALDANE, J. S., 1920, *Methods of air analysis* (London: Charles Griffin).
HILL, D. W., 1960, *Gas chromatography* 1960 (Ed. R. P. W. Scott) (London. Butterworth), p. 344.
JAHN, R. E., 1955, *The Analyst*, **80,** 700.
LEWIS, G. N., and RANDALL, M., 1923, *Thermodynamics and the Free energy of chemical substances* (New York: McGraw-Hill).
LLOYD, B., 1958, *J. Physiol.*, **143,** 5P.
LUFT, K. F., 1943, *Z. Tech. Phys.*, **24,** 97.
MAAS, O., and GEDDES, A. C., 1937, *Phil. Trans A*, **236,** 303.
MICHELS, A., SCHAMP, H. W., and DE GRAAFF, W., 1954, *Physica*, **20,** 1209.
NUNN, J. F., 1958, *Brit. J. Anaesth.*, **30,** 254.
SCHOLANDER, P. F., 1947, *J. Biol. Chem.*, **167,** 235.
TIMMS, D. G., KONRATH, H. J., and CHINSIDE, R. C., 1958, *The Analyst*, **83,** 600.

Four-probe resistivity measurements on small circular specimens

by D. E. VAUGHAN, B.Sc., Ph.D., A.R.C.S., Physics Branch, Royal Military College of Science, Shrivenham, Swindon, Wiltshire

MS. received 22*nd March* 1961

Abstract

General formulae are presented for square and linear probe arrays on circular specimens which permit the calculation of sheet resistivity from voltage and current measurements made with the probes in any position. From these formulae it is deduced that the square probe array is preferable on small specimens on the grounds of accuracy, as well as from spatial considerations.

Introduction

In the course of work on the properties of irradiated germanium it became necessary to make four-probe resistivity measurements on thin circular specimens of small diameter. On account of the small diameter it was clearly advantageous to use a square rather than a linear probe configuration. Formulae have been published for linear probe configurations on square and circular specimens by Smits (1958), and for square probe configurations on square specimens by Keywell and Dorosheski (1960), but not, apparently, for square configurations on circular specimens. In addition, no reference has been made to the effect on the formulae of displacement of the centre of the probe array from the geometrical centre of the specimen.

Formulae for the square probe array

From elementary theory it can be shown that for a specimen of uniform thickness and resistivity:

$$\rho_s = \left(\frac{V}{I}\right)\frac{2\pi}{\ln 2S} \tag{1}$$

where ρ_s is the sheet resistivity, V is the potential difference between the voltage probes, I is the current flowing between the current probes, and S is the correcting factor applicable to square probe arrays on finite specimens. The derivation of formula (1) is given in the appendix.

For a square probe array of side $2t$ on a circular specimen of radius r, where the centre of the probe array is displaced a distance αt in a direction parallel to the current dipole and a distance βt perpendicular to the current dipole, it can be shown that

$$S_{\alpha,\beta}{}^2 = \frac{\{ADR^4 - (A + D - 8)R^2 + 1\}\{BCR^4 - (B + C - 8)R^2 + 1\}}{\{ACR^4 - (A + C - 4)R^2 + 1\}\{BDR^4 - (B + D - 4)R^2 + 1\}} \tag{2}$$

where $A = (\alpha - 1)^2 + (\beta + 1)^2$, $C = (\alpha - 1)^2 + (\beta - 1)^2$, $B = (\alpha + 1)^2 + (\beta + 1)^2$, $D = (\alpha + 1)^2 + (\beta - 1)^2$

and $R = t/r$.

When the total displacement is perpendicular to the current dipole $\alpha = 0$, whence $A = B = 1 + (1 + \beta)^2$ and $C = D = 1 + (1 - \beta)^2$. Substitution of these values in (2) gives

$$S_{0,\beta} = \frac{\{ADR^4 - (A + D - 8)R^2 + 1\}}{\{ADR^4 - (A + D - 4)R^2 + 1\}} \tag{3}$$

$$= 1 + \frac{4R^2}{(\beta^2R^2 - 1)^2 + 4R^4}. \tag{4}$$

When the total displacement is parallel to the current dipole $\beta = 0$, whence $A = C$ and $B = D$, but in this case substitution for A and B in terms of α effects no simplification.

From simple geometrical considerations it can be shown that when the current probes lie at the circumference of the specimen

$$\frac{1}{R^2} = 1 + (1 + \beta)^2. \tag{5}$$

This, when substituted in (3), reduces it to $S = 2$. Since this result holds irrespective of the position of the voltage probes, provided the square configuration still obtains, it must be true when these, also, are at the circumference. This result is then identical with that obtained by Van der Pauw (1958) using a different approach.

Formulae for the linear probe array

The formula corresponding to (1) is as follows:

$$\rho_s = \left(\frac{V}{I}\right)\frac{2\pi}{\ln 4L} \tag{6}$$

where ρ_s, V and I are defined as before and L is the correcting factor applicable to linear probe arrays on finite specimens. Formula (6) is also derived in the appendix.

For a linear probe array with inter-probe spacing $2t$ on a circular specimen of radius r, where the mid-point of the probe array is displaced a distance t in a direction parallel to the current dipole and a distance t perpendicular to the current dipole, the formula corresponding to (2) is as follows:

$$L_{\alpha,\beta}{}^2 = \frac{\{EHR^4 - (E + H - 16)R^2 + 1\}\{FGR^4 - (F + G - 16)R^2 + 1\}}{\{EGR^4 - (E + G - 4)R^2 + 1\}\{FHR^4 - (F + H - 4)R^2 + 1\}} \tag{7}$$

where $E = (3 + \alpha)^2 + \beta^2$ $\quad G = (1 + \alpha)^2 + \beta^2$
$F = (3 - \alpha)^2 + \beta^2$ $\quad H = (1 - \alpha)^2 + \beta^2$
and $R = t/r$.

When the total displacement is perpendicular to the

current dipole $\alpha = 0$, whence $E = F = 9 + \beta^2$ and $G = H = 1 + \beta^2$. Substitution of these values in (7) gives

$$L_{0,\beta} = \frac{\{EGR^4 - (E + G - 16)R^2 + 1\}}{\{EGR^4 - (E + G - 4)R^2 + 1\}} \tag{8}$$

$$= 1 + \frac{12R^2}{\{(\beta^2 + 3)R^2 - 1\}^2 + 4\beta^2R^2}. \tag{9}$$

When the total displacement is parallel to the current dipole $\beta = 0$, and, again, no simplification is possible.

The formula corresponding to (5) is as follows:

$$\frac{1}{R^2} = \beta^2 + 9 \tag{10}$$

which, when substituted in (9), reduces it to $L = 4$. This result is identical with that which can be derived from basic theory (see Appendix).

Discussion

In this discussion it will be assumed that the t/r ratio lies between 1/3 and 1/5 and that the minimum displacement factor is 1/20. These values are typical of the problem which initiated this work.

Although corresponding formulae for linear and square probe arrays are very similar in form, differentiation of formulae (4) and (9) shows that, if the displacement factor is regarded as an error, then the error induced in $L_{0,\beta}$ is at least four times the corresponding error induced in $S_{0,\beta}$. A precise calculation, using simultaneous non-zero values for α and β is difficult to interpret in the general case, but it is obvious that for small values of α and β, when the $\alpha\beta$ product terms are negligible, the errors are additive. These results, coupled with the fact that the forms of variation of L and S with respect to both α and β are such that the errors in them increase rapidly with increasing displacement factor, for a given displacement error, make the square probe array preferable to the linear probe array. Also, the obvious fact that the square configuration can be used on smaller specimens carries with it the cognate advantage that, for a given value of the t/r ratio, the probes are farther from the edges of the specimen where non-uniformities, if they existed, would introduce errors.

From formulae (3) and (8) it is obvious that the interchange of the current and the voltage probes has no effect on the values of $L_{0,\beta}$ and $S_{0,\beta}$. This can be used as a qualitative test for the uniformity of the specimen. This test is most useful when the t/r ratio is small. It is hoped to publish a paper on the quantitative aspects of such measurements at a later date.

It is interesting to note the contrast between the relative simplicity of the expressions for $S_{0,\beta}$ and $L_{0,\beta}$ as given in formulae (3) and (8) and the complexity of the corresponding formulae for $S_{\alpha,0}$ and $L_{\alpha,0}$. This can be explained by the element of symmetry existing in the former translation which does not occur in the latter.

All the formulae quoted in this paper are based upon a two-dimensional model so that it is necessary to know, in practical terms, the maximum permissible thickness of specimen to which they are still applicable. Data published by Smits (1958) show that, if w is the thickness of the specimen and $2t$ is the inter-probe spacing, two-dimensional formulae are accurate to within 0·25% if the ratio w/t is less than unity.

Appendix

Harnwell (1938) shows that for a single current source in an infinite sheet the potential at a point a distance d from the source is given by

$$\phi - \phi_0 = -\frac{I}{2\pi}\rho_s \ln d \tag{11}$$

where ϕ is the potential in question, ϕ_0 is some arbitrary reference potential, ρ_s the sheet resistivity and I the current flowing. From this it follows that the potential at a point distance d_1 and d_2 from the poles of the current dipole is given by

$$\phi - \phi_0 = -\frac{I}{2\pi}\rho_s \ln \frac{d_2}{d_1}. \tag{12}$$

Therefore the potential difference between two points at

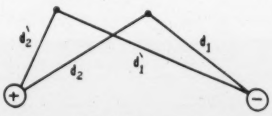


Fig. 1. Simple dipole configuration.

distances d_1 and d_2, and d_1' and d_2' from a current dipole (see Fig. 1) is given by the following general expression:

$$\Delta\phi = \frac{I\rho_s}{2\pi} \ln \frac{d_2 d_1'}{d_1 d_2'} \tag{13}$$

where $\Delta\phi$ is identical with the V used in formulae (1) and (6).

For linear probe arrays (see Fig. 2) with equal inter-

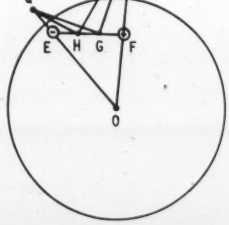


Fig. 2. Linear probe configuration.

probe spacing $d_1' = d_2 = 2d_1 = 2d_2'$. Substituting these values in (13) gives

$$\rho_s = \left(\frac{V}{I}\right)\frac{2\pi}{\ln 4}. \tag{14}$$

This is identical with formula (6) applied to an infinite sheet, since the substitution of $t/r = 0$ in formula (7) gives $L = 1$.

For a circular finite sheet an additional image dipole PQ is introduced to maintain the necessary boundary conditions. This adds an additional term to formula (13) to give

$$\Delta\phi = \frac{I\rho_s}{2\pi}\left(\ln 4 + \ln \frac{PH \cdot QG}{PG \cdot QH}\right). \tag{15}$$

This formula can be written in the form

$$\rho_s = \left(\frac{V}{I}\right)\frac{2\pi}{\ln 4L} \tag{16}$$

which is identical with formula (6).

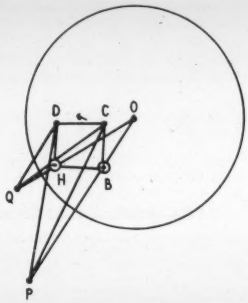


Fig. 3. Square probe configuration.

For the square probe array (see Fig. 3) $d_1' = d_2 = \sqrt{2}d_1 = \sqrt{2}d_2'$. Substituting these values in (13) gives

$$\rho_s = \left(\frac{V}{I}\right)\frac{2\pi}{\ln 2}. \tag{17}$$

This is identical with formula (1) applied to an infinite sheet, since the substitution of $t/r = 0$ in formula (4) gives $S = 1$.

Introducing the image dipole as before gives

$$\Delta\phi = \frac{I\rho_s}{2\pi}\left(\ln 2 + \ln\frac{PD.QC}{PC.QD}\right). \tag{18}$$

This formula can be written in the form

$$\rho_s = \left(\frac{V}{I}\right)\frac{2\pi}{\ln 2S} \tag{19}$$

which is identical with formula (1).

References

HARNWELL, G. P., 1938, *Principles of Electricity and Electromagnetism.* Chapter I (New York: McGraw-Hill).

KEYWELL, F., and DOROSHESKI, G., 1960, *Rev. Sci. Instrum.*, **31**, 833.

SMITS, F. M., 1958, *Bell Syst. Tech. J.*, **37**, 711.

VAN DER PAUW, L. J., 1958, *Philips Res. Rep.*, **13**, 1.

Dielectric properties of ammonium dihydrogen phosphate at very high frequencies

by E. RUSHTON, B.Sc., National Physical Laboratory, Teddington, Middlesex

MS. received 11*th April* 1961

Abstract

Ammonium dihydrogen phosphate shows no detectable dielectric dispersion between 10 *kc/s and* 36 *Gc/s. Determinations of the permittivity* ϵ *at a frequency of* 36 *Gc/s are described: the values obtained are* $\epsilon_{\parallel} = 14{\cdot}0$ *at* 21° C *and* $\epsilon_{\perp} = 57{\cdot}1$ *at* 21·5° C, *where* $\epsilon_{\parallel}$ *and* $\epsilon_{\perp}$ *are the values of permittivity when the applied electric field is parallel with or perpendicular to the optical axis. These values are not measurably different from those obtained at lower frequencies by the author (at* 1 *Mc/s) and by other workers (at* 10 *kc/s).*

Introduction

AMMONIUM dihydrogen phosphate has been used recently by Froome and Bradsell (1960) as an electro-optical shutter to modulate a beam of light at microwave frequencies, and for this work a knowledge of the dielectric properties of the material at centimetre wavelengths is needed. Published information about the dielectric properties at very high frequencies is scanty. Measurements of the permittivity have been made by Bechmann *et al.* (1957) at a frequency of 10 kc/s and a temperature of 18·5° C, and the values reported by them are as follows: $\epsilon_{\parallel} = 14{\cdot}0 \pm 0{\cdot}2$ and $\epsilon_{\perp} = 57{\cdot}6 \pm 0{\cdot}5$, where $\epsilon_{\parallel}$ denotes the value of permittivity when the applied electric field is parallel with the optical axis and $\epsilon_{\perp}$ the value when the electric field is perpendicular to the optical axis. Published figures for higher frequencies (von Hippel 1954) indicate that at 25° C the value of $\epsilon_{\parallel}$ is 13·7 at 10 Gc/s and the value of $\epsilon_{\perp}$ is 55·9 at 300 Mc/s, but since the values at 10 kc/s ($\epsilon_{\parallel} = 15{\cdot}4$ and $\epsilon_{\perp} = 55{\cdot}9$) differed appreciably from Bechmann's values, it seemed desirable to make a fresh determination of the values at very high frequencies.

Method of measurement

The highest frequency at which measurements can be made at present at the National Physical Laboratory is 36 Gc/s, corresponding to a wavelength of about 8 mm. The method of measurement is shown in Fig. 1. The source is a klystron

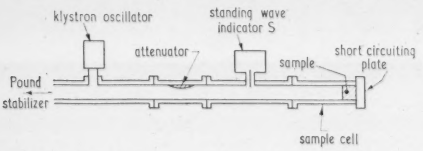


Fig. 1. Apparatus for measurements at 36 Gc/s by standing wave method.

generating power at the required frequency, and the material under examination is fitted at the end of the sample cell, which forms part of the wave-guide system. The standing wave pattern is examined by means of the standing wave indicator S consisting of a travelling probe projecting a short distance into the wave-guide: the r.f. output from the probe is rectified and monitored by means of a d.c. galvanometer. When no material is in the sample cell, the travelling wave is reflected at the short-circuiting plate terminating the sample cell and the standing wave resulting from the interaction of the incident and reflected waves has minimum values or nodes which are nearly zero. The nodes are detected by the standing wave indicator, and the distance x of the first node from the short-circuiting plate at the end of the sample cell is measured. When the sample of material is fitted at the end of the sample cell there is some absorption of energy, the amount depending on the material, and the standing wave pattern is altered since the amplitude of the reflected wave is reduced because of absorption in the material and the standing wave minimum is no longer zero. Further, since the permittivity of the material is greater than unity, the positions of the nodes are shifted, and the distance of the first node from the short-circuiting plate is increased to a value x_s. It can be shown that the permittivity ϵ is given by

$$\epsilon = \frac{\lambda_c^{-2} + (\beta_2/2\pi)^2}{\lambda_c^{-2} + \lambda_g^{-2}} \qquad (1)$$

where λ_c = cut-off wavelength for the guide, λ_g = wavelength in the guide and β_2 is derived from the expression

$$\frac{\tan \beta_2 d}{\beta_2 d} = \frac{-\lambda_g \tan (2\pi x_0/\lambda_g)}{2\pi d} \qquad (2)$$

where d = thickness of specimen and x_0 = distance of the first node from the surface of the specimen $= x_s - d$. Since the quantities λ_c, λ_g, d and x_0 can all be derived from the measurements, $\beta_2 d$, and hence β_2, is obtained from Eqn (2). There are, however, an infinite number of solutions to Eqn (2), giving different values of $\beta_2 d$, and the procedure adopted was to measure two samples of differing thickness and so derive a common value for β_2, which was then substituted in Eqn (1) to give the value for permittivity.

In the H_{10} wave-guide system used, there are two components of the magnetic field H, one longitudinal and the other transverse, and also a transverse electric field E (Fig. 2). To determine $\epsilon_{\parallel}$ or $\epsilon_{\perp}$, samples of ammonium dihydrogen phosphate were cut so that when they were inserted in the wave-guide the transverse electric field in the guide was parallel (Fig. 2(*a*)) or perpendicular (Fig. 2(*b*)) to the optical axis. The pieces had to be accurately cut so that they were a close fit in the wave-guide, and this operation involving

cutting flat pieces of the required dimensions ($a = 7{\cdot}26$ mm, $b = 3{\cdot}58$ mm and $d = 2$ to 3 mm) from the rather brittle material was skilfully carried out by Mr. J. Conway (late of

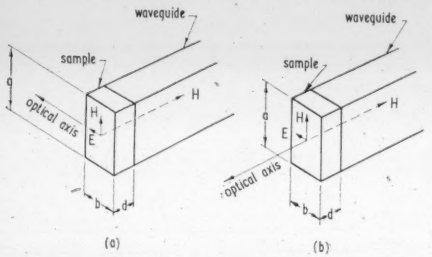


Fig.-2. Sample in wave-guide, (*a*) electric field E parallel to optical axis, (*b*) electric field E perpendicular to optical axis.

the Standards Division Workshop). When satisfactory pieces had been prepared, all the five faces in contact with the sides and the end-plate of the sample cell were coated with evaporated silver films: this was done because intimate electrical contact between the sample faces and the metallic sides and end-plate is essential when measuring materials of relatively high permittivity.

Measurements of permittivity

Measurements on two samples of ammonium dihydrogen phosphate of differing thickness were made at a frequency of 36 Gc/s with the electric field parallel with the optical axis. The value found for the permittivity for each sample was $\epsilon_{\parallel} = 14{\cdot}0$ at 21° C. The estimated limit of error was $\pm 0{\cdot}3$. It was a relatively simple matter to measure the permittivity at a lower frequency of 1 Mc/s and for this measurement a larger piece, 19 mm diameter and 1·8 mm thick, was cut and the permittivity determined by means of the Hartshorn–Ward (1936) method. The value obtained at 1 Mc/s and at 21° C, was $\epsilon_{\parallel} = 14{\cdot}0$, the estimated limit of error being $\pm 0{\cdot}2$.

Thus the value of $\epsilon_{\parallel}$ is, within the experimental error, the same at 36 Gc/s as at 1 Mc/s and it is also the same as that reported by Bechmann at 10 kc/s.

Measurements were made using the same technique at 36 Gc/s on two samples with the electric field perpendicular to the optical axis. The values of $\epsilon_{\perp}$ obtained were 56·7 for one sample and 57·4 for the other, but here the limits of error were estimated to be ± 1. Fortunately, more accurate measurements of $\epsilon_{\perp}$ at 36 Gc/s could be made by an alternative method using a cylindrical cavity resonator (Horner *et al.* 1946). This method is well known, and all that needs to be said here is that the permittivity can be determined from measurements of the resonant length and Q-factor of the cavity made with the cavity empty and then with a sample of material introduced into the cavity. The cylindrical cavity used for the measurements was 19 mm diameter, and a sample of approximately this diameter was cut to fit in the cavity. In this case, however, a close fit in the cavity was not essential since the transverse circumferential electric field is zero at the walls of the cavity. It was a relatively simple matter to cut a flat disk of the material of the required dimensions, such that the optical axis was perpendicular to the electric field in the cavity. (A sample of the same size but with the optical axis parallel to the electric field could not be cut from the material available.) The value obtained by the cylindrical cavity method at 36 Gc/s was $\epsilon_{\perp} = 57{\cdot}1$ at 21·5° C. The estimated limit of error was $\pm 0{\cdot}6$.

At a lower frequency of 1 Mc/s the result obtained was $\epsilon_{\perp} = 57{\cdot}2$ at 21·5° C, the estimated limit of error being $\pm 0{\cdot}6$. Here, again, the values at 36 Gc/s and 1 Mc/s are in good agreement, and (taking account of the temperature coefficient of permittivity $d\epsilon/\epsilon dt = -25 \times 10^{-4}$ per deg C) they are in good agreement too with Bechmann's value of $\epsilon_{\perp} = 57{\cdot}6 \pm 0{\cdot}5$ at 10 kc/s and 18·5° C.

The values of loss tangent at 36 Gc/s in the two cases are of interest: the value of $\tan \delta_{\parallel}$ (i.e. with the field parallel to the optical axis) was 0·025 and the value of $\tan \delta_{\perp}$ was 0·07.

Acknowledgment

The work described has been carried out as part of the research programme of the National Physical Laboratory and is published by permission of the Director of the Laboratory.

References

BECHMANN, R., PARSONS, P. L., HOLLINS, G. T., and AYERS, S., 1957, *Piezoelectricity. Selected engineering reports of the Post Office Research Station* (*Report No.* 8), (London: H.M.S.O.), p. 241.

FROOME, K. D., and BRADSELL, R. H., 1960, British Patent Application No. 32675/60.

HARTSHORN, L., and WARD, W. H., 1936, *J. Instn Elect. Engrs*, **79**, 597.

VON HIPPEL, A. R. (Ed.), 1954, *Dielectric Materials and Applications* (London and New York: Wiley), p. 301.

HORNER, F., TAYLOR, T. A., DUNSMUIR, R., LAMB, J., and JACKSON, W., 1946, *J. Instn Elect. Engrs*, **93**, 53.

Time lags in the electrical breakdown of glass immersed in water

by M. N. AZAM, B.Sc., Ph.D., and H. DICKINSON, B.Sc.Tech., A.M.I.E.E., Department of Engineering, University College of Swansea, University of Wales

MS. received 6th August 1960, *in revised form* 13*th March* 1961

Abstract

When measuring the electric breakdown strength of cover glass immersed in deionized water, time lags to breakdown were observed. The mean statistical time lag was (12 ± 1) *μs. The breakdown strength between spherical electrodes was found to be* $(11{\cdot}4 \pm 1) \times 10^6$ *v/cm.*

1. Introduction

In recent years some interest has been taken in the study of time lags in the electrical breakdown of solid dielectrics. Whitehead (1951), in a review of the information, suggested them to be less than 10^{-7} second. Some of the work done by Kawamura and associates (1954) on potassium chloride crystals and glass, and Konorova (1957) on mica and glass, is in agreement with this view. The work of Inuishi *et al.* (1952) and Cooper and Grossart (1956) on potassium chloride crystals, and also of Kawamura *et al.* (1954) on mica, suggests rather longer time lags (10^{-5} second). This discrepancy led the authors to investigate this effect in the breakdown of thin specimens of cover glass of known composition. The specimen thickness was about 0·01 cm in all cases.

In order to overcome the effects of spurious external discharges in the breakdown experiments, several investigators have used either specimens with ground recesses or specimens in the form of bulbs blown on the end of glass tubes. In the latter case internal and external liquid electrodes were used, whilst recessed specimens have had pressure contact electrodes with dielectric guard rings. To avoid these relatively difficult preparations it was decided to make measurements on microscope cover glasses, immersed in a medium with high dielectric constant and resistivity. After preliminary tests with a variety of media, deionized water was finally chosen. The requirements of suitable dielectric immersion media have been fully discussed by Standring (1941). The specimens were first examined for any defects and then measured for uniformity of thickness. Only those within $\pm 0{\cdot}000\,25$ cm of a given thickness were accepted. They were then immersed in trichlorethylene for one minute, placed in acetone for one minute, and finally washed with distilled water and dried in a desiccator.

2. Experimental method

The measurements were made using a glass cell 15 cm × 13 cm × 18 cm, fitted with two adjustable 1 cm diameter bronze spherical electrodes. This apparatus was carefully cleaned between tests. The glass specimens were placed between the electrodes, and the spheres were adjusted until there was sufficient pressure to hold the specimens in place. The cell was then filled with deionized water.

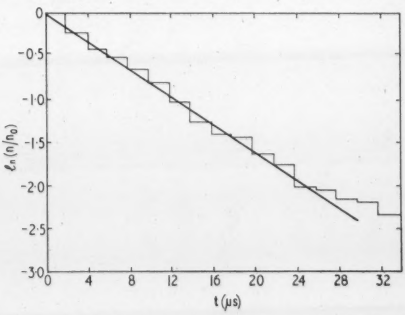
The method of measurement consisted of applying to each specimen a series of 0·1/1500 μs pulses until breakdown occurred. The amplitude of each successive impulse exceeded that of the previous one by about 3%. A capacitive potential divider enabled the waveform to be displayed and photographed on a calibrated cathode-ray oscilloscope. The voltages were constant, within 1%, for 35 μs from the peak of the impulse. This time is considerably longer than the time lags observed in the breakdown measurements for most of the specimens. All the measurements were made at room temperature.

3. Results

The experiments with immersion media of dielectric constant of 35 or less gave breakdown paths, through the glass specimens, at some distance away from the axis of the electrode system.

When water was used as the immersion medium many of the glass specimens shattered on breakdown, but those that remained intact were carefully examined. As far as could be determined by visual examination of the specimens and of the electrodes the breakdown path was always on the axis of the electrodes.

The observed distribution of time lags, for 97 specimens,



Distribution of time lags in the breakdown of glass immersed in deionized water (97 specimens at room temperature).

is shown in the Figure. This distribution can be represented by the equation

$$n = n_0 \exp(-t/\tau)$$

where n is the number of specimens breaking down with time lags exceeding t μs; n_0 is the total number of specimens tested; τ is the mean statistical time lag. The mean statistical time lag of the distribution shown in the Figure was found to be (12 ± 1) μs.

The apparent electric strength of the cover glasses under the described experimental conditions was $(11{\cdot}4 \pm 1) \times 10^6$ v/cm.

4. Discussion

The position of the breakdown paths when the specimens were immersed in deionized water and the high value of the electric breakdown strength obtained, strongly suggests that the measured breakdown strength is characteristic of the specimen itself. The use of media with dielectric constants exceeding that of water may enable the validity of the experimental procedure for measuring the intrinsic strength of glass, as discussed above, to be established. Some organic compounds with dielectric constants of 130 or more are available and may prove suitable for use as immersion media.

The present investigation shows that the practice, often used in measuring intrinsic electric breakdown strength of glass, of applying a single linearly rising voltage pulse of a few microseconds duration may yield values that are appreciably higher than the true value, because of the possible overshoot of the voltage caused by the relatively long time lags. For example, Keller (1951) and Vermeer (1956) obtained results with glass that showed higher electric breakdown strengths with faster rising voltage waveforms.

Acknowledgments

We are grateful to Professor W. Fishwick for the facilities provided in his department. We also wish to thank our colleague Dr. W. G. Townsend for valuable suggestions and discussions.

References

Cooper, R., and Grossart, D. T., 1956, *Proc. Phys. Soc. B*, **69**, 1351.

Iniushi, Y., Yamanaka, C., and Suita, T., 1952, *Technol. Rep. Osaka Univ.*, **2** (31).

Kawamura, H., Ohkura, H., and Kikuchi, T., 1954, *J. Phys. Soc. Japan*, **2**, 541.

Keller, K. J., 1951, *Physica*, **17**, 511.

Konorova, E. A., 1957, *J. Expt. Theoret. Phys. USSR*, **32**, 603.

Standring, W. G., 1941, *J. Instn Elect. Engrs*, Part II, **88**, 360.

Vermeer, J., 1956, *Physica*, **22**, 1247, 1274.

Whitehead, S., 1951, *Dielectric Breakdown of Solids* (Oxford: University Press).

Method of determining the specimen–surface displacement of an X-ray diffractometer

by B. W. DELF, B.Sc., Grad.Inst.P., Viriamu Jones Laboratory, University College, Cardiff

MS. received 20*th January* 1961, *in revised form* 23*rd March* 1961

Abstract

A simple method of determining the specimen–surface displacement of an x-ray diffractometer is described. The position of the specimen surface is compared with the position of a glass slide of known specimen–surface displacement by means of a travelling microscope. The displacement of a specimen is found to 0·006 *mm which corresponds to a maximum angular error of* 0·004° *in* 2θ *for the Philips diffractometer used, radius* 17 *cm. The displacement is shown to be critically dependent on the method of locking the specimen holder. This is overcome by performing the various steps of the locking procedure in a certain fixed order. The effect of movement of the specimen holder shaft on the displacement is discussed.*

Introduction

In diffraction patterns obtained with an x-ray diffractometer, the positions of both the peaks and the centroids of the diffraction maxima are displaced from their true positions by various instrumental and diffraction effects. Most of these effects can be corrected for, both theoretically and practically. The major trouble, from an experimental point of view, is caused by the displacement of the centre of the surface of the specimen from the axis of rotation of the diffractometer (S–S displacement). The magnitude of the effect has been given by Wilson (1950).

The effect has been dealt with in a number of ways, the most common of which is the use of a calibrating substance (for example, Swanson, Gilfrich and Cook 1957). In the determination of lattice parameters extrapolation against $\cos^2\theta$ to $\theta = 90°$, where the effect disappears, has been used by Coyle and Garrod, while another approach is due to Tournarie (both methods have been briefly described by Parrish 1960). None of these methods is completely satisfactory for the precise determination of lattice parameters by means of centroids.

The S–S displacement error is caused by a combination of two factors. The first is that the axis of rotation of the diffractometer is not necessarily either coincident with the banking surface against which the specimen mount is held, or parallel to it. Secondly, the surface of the specimen may not be in the plane of the top surface of the specimen mount.

A method is required, therefore, for determining both the slope of the surface relative to the axis of rotation, and the S–S displacement of the specimen. In the method which has been devised, the position of the surface of the specimen is compared with the position of a glass slide, the S–S displacement of which has been accurately measured. This enables the appropriate correction to be made to the position of each peak or centroid.

Principle of method

The method used is in essence extremely simple and requires little in the way of apparatus; a travelling microscope, fitted with an objective lens of short focal length (one of 4 mm was used for the work described) is all that is necessary.

A thin glass rod is mounted on the banking surface with its long axis along the axis of the diffractometer (Fig. 1). The position of the rod is adjusted until, on turning the specimen holder through 180°, there is no movement of an easily viewed and identified point on the rod. This means, of course, that the axis of the diffractometer, which is horizontal in the instrument used (Fig. 1), passes through the point in question.

A travelling microscope fitted with the objective lens of short focal length is mounted above the rod in such a way that it can be used to measure distances in the vertical direction. The microscope is then focused on the point on the rod which remained stationary on rotation.

The glass rod is removed from the banking surface of the specimen holder and a glass slide put in its place; the specimen holder is adjusted so that the plane of the glass slide is horizontal. The microscope is then focused on the top surface of the glass slide, and hence the perpendicular distance between the diffractometer axis and the top surface of the glass slide at a particular point on the diffractometer axis is found. This is the S–S displacement of the glass slide at this point. The particular point on the diffractometer axis to which this measurement refers is defined by the distance between the point on the glass rod which remained stationary on rotation and the point B of Fig. 1, which is the front edge of the banking surface. The above procedure is repeated at various distances from B.

The S–S displacement of a specimen in general varies with distance from B, and hence a mean value has to be taken. Provided that the variation is linear, as it has been found to be, this mean value is the S–S displacement of the centre of the irradiated area of the specimen. (In the diffractometer used this point is 0·60 cm from B.) The S–S displacement of the glass slide at this point can readily be determined and this enables the mean S–S displacement of a specimen to be found as follows.

The glass slide is mounted horizontally in the specimen holder and the travelling microscope, mounted as before, is focused on the point on the top surface of the glass slide, which is 0·60 cm from B. The slide is then removed, a specimen mount containing a specimen put in its place, and the perpendicular distance between the point on the glass slide and the specimen can be found. The S–S displacement of the specimen at the point which is the centre of the irradiated part of the specimen can then be calculated.

Practical details of procedure

The thin glass rod, which is about 0·01 mm in diameter (A in Fig. 1), is mounted by means of 'Plasticine' on the

banking surface of the specimen holder, with its axis approximately parallel to the axis of rotation of the diffractometer. The specimen holder is unlocked from the gears to enable it to be rotated freely through 360°.

The rod is then viewed end-on by means of the microscope

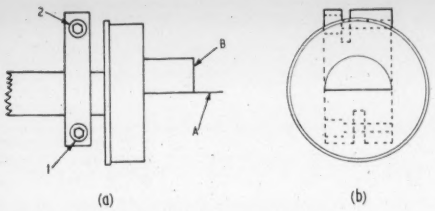


Fig. 1. The specimen holder viewed (*a*) perpendicular to the diffractometer axis and (*b*) along the axis.
A, thin glass rod; B, front edge of the banking surface; 1, 2, Allen screws.

fitted with the 4 mm objective lens, and with the specimen holder in the position shown in Fig. 1. The easily viewed point on the rod to be used is on the top surface of the end of the rod, when the specimen holder is in its original position. The position of the glass rod is adjusted by pressing the 'Plasticine' until there is no visible movement of this point on turning the specimen through 180°. The knack of doing this is easily acquired after three or four attempts.

The specimen holder is then locked with the plane of the banking surface approximately horizontal, as in Fig. 1, and the end of the glass rod observed from above with the travelling microscope fitted with the same objective lens as before. The microscope is focused on the top surface of the end of the rod and the position noted by means of a large divided circular scale, which is attached to the micrometer screw of the microscope, and which is later calibrated against a divided millimetre. The distance of the end of the rod from the edge B (Fig. 1) of the banking surface is then found and the rod removed.

A glass slide is then placed in the specimen holder and the microscope refocused on its top surface and the S–S displacement of the slide calculated. This is repeated at various distances from the edge of the banking surface in order to determine the slope of the glass slide relative to the axis of the diffractometer and hence the S–S displacement at the point corresponding to the centre of the irradiated area of the specimen. Using this result the mean S–S displacement of a specimen can be determined as described above.

The greatest experimental difficulty in this method is in locating the top surfaces of the slide and the rod when viewing from above. Several techniques for making these surfaces visible were tried, including grinding, etching and the use of a microcrystalline powder, but all were found to have disadvantages. It was eventually found that focusing on very small dust particles (<0·001 mm) is the best means of locating these surfaces.

For the slide this procedure is not too critical, because the slide is only being used as an intermediary and, provided the same sort of particles are used in both parts of the procedure, no appreciable error should result. For the glass rod, however, focusing on the top surface is quite critical. In this case the source of light used was adjusted to illuminate the top edge of the rod and the position of this compared with the position of the top surface given by focusing on the dust particles. In no case was a difference greater than 0·002 mm found, and the mean position was taken as that of the top surface of the rod.

It is important that the microscope should be focused on the slide surface at the same point in the field of view as was the end of the rod, as the plane of the slide may not be perpendicular to the axis of the microscope. The field of view of the microscope is about 0·5 mm in diameter, and a mis-setting of 1° can produce an error of 0·01 mm.

The time taken to calibrate a slide is a matter of hours. For example, the results for each straight line of Fig. 2 took about 3 hours to obtain. Once these results have been obtained, however, the S–S displacement of a specimen may be found in 5 minutes.

Accuracy

The depth of focus of the optical system of the microscope is about 0·001 mm, and this produces an uncertainty in the comparison between the positions of the glass slide and the rod of 0·002 mm. This, coupled with the uncertainty of 0·002 mm involved in setting the glass rod, produces an uncertainty in the S–S displacement of the glass slide at any point of about 0·003 mm.

The S–S displacement of the slide at the point corresponding to the centre of a specimen is obtained from a number of readings, but the resultant increase in accuracy is counterbalanced to some extent by the fact that, because of the construction of the diffractometer, it is impossible to obtain measurements nearer than 0·3 cm to the centre of the specimen when using a normal microscope objective lens. In order to overcome this difficulty a special objective lens with a barrel diameter of only 0·9 cm was purchased from J. Swift and Son. With this lens measurements can be taken right through the point corresponding to the centre of specimen C (Fig. 2) and about 0·15 cm beyond. The

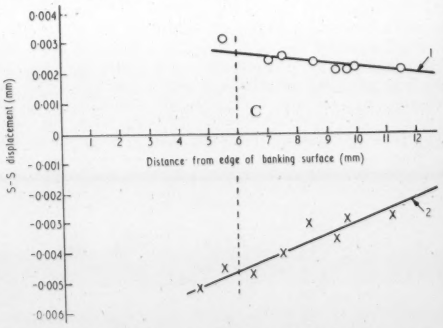


Fig. 2. The variation of S–S displacement of a glass slide with distance from the edge of the banking surface, before, curve 1, and after, curve 2, moving the specimen holder shaft.
C corresponds to the centre of a diffractometer specimen.

uncertainty in the S–S displacement of the slide at the centre of the specimen is therefore estimated at 0·002 mm, which can produce an uncertainty in the S–S displacement of the centre of a specimen of 0·003 mm.

There are, however, several factors with regard to the preparation of the specimen surface, and in particular with regard to the specimen being parallel to the slide, which can

introduce additional uncertainties. These are reduced to a certain extent by taking measurements of the difference in the positions of the slide and the surface at varying distances from the holder. The uncertainty in the S–S displacement of the centre of the specimen will still be increased, but is unlikely to be greater than 0·006 mm. This corresponds to a maximum error of 0·004° in 2θ with a diffractometer radius of 17 cm.

Instrumental limitations

The most important result obtained by the above method is probably the knowledge of the limitations of the diffractometer, as far as the S–S displacement is concerned, rather than the measurement of the effect.

In the Philips diffractometer type P.W1 1050/30 used in this work, the specimen holder has to be unlocked from the 2 : 1 gears in order to measure the S–S displacement and then locked again before any x-ray diffraction work is undertaken. It was found that under normal circumstances a very large shift of the specimen surface could be produced on relocking the specimen holder. The largest such movement observed was approximately 0·07 mm, and the average was about 0·05 mm.

At first it was thought that the only approach would be to assume that the movement was about the axis and measure the magnitude of the shift, but further investigation of the locking mechanism led to a method of eliminating this effect.

The specimen holder shaft is locked to the 2 : 1 gears by means of a clamp. This clamp is tightened on to the shaft by an Allen screw (1), Fig. 1, and to a locating pin by a second screw (2). The whole system is held firmly by a shaft which screws in from the rear. In order to allow the specimen holder to be rotated through 180° this shaft has to be removed and screw 1 unlocked, together with the two screws which control the fine and coarse adjustment of the 2 : 1 ratio. Locking and unlocking this system in any order always produced a shift of the specimen surface.

It was then found that if screw 2 were unlocked, it was possible in certain circumstances to lock the specimen shaft without any movement of the specimen surface. The procedure adopted was first to lock the screws controlling the 2 : 1 ratio in any order, then lock screw 1, next screw 2 and finally to replace the shaft. It was found quite essential that the last three be done in the order stated. The tightening of the shaft was the factor which was capable of producing the greatest movement, and this was done with some care.

It would not appear to be certain that the above procedure would produce no movement of the specimen surface if used on another diffractometer, even if the instrument were of the same design. It is advisable that the effect on the position of the specimen surface of locking the various screws should be investigated for any other instrument, and a satisfactory locking procedure found. For the present instrument, however, this procedure has been found to be entirely satisfactory. The locking procedure has been used in the course of some fifty measurements and no appreciable movement noted. On occasions the surface was seen to move while screw 1 was being tightened, but it invariably returned to its original position when the operation was complete.

A second and equally important instrumental limitation concerns the specimen-holder shaft. It has been found that if the shaft is taken out and replaced, or even moved forward and replaced, the position of the glass slide with relation to the axis and its inclination to it can be changed.

As a typical example, Fig. 2 shows the change obtained on moving the shaft out and replacing it. This effect is obviously not important as long as the holder is not moved. If it has to be moved, or is moved inadvertently, the calibration of the glass slide should be checked and if necessary the calibration repeated.

These limitations must be kept in mind while determining the S–S displacement, but provided this is done the S–S displacement of the centre of a specimen may be determined to the accuracy suggested above. This is shown by the close approximation of the experimental points to straight lines (Fig. 2).

The specimen holder and the banking surface of the instrument were potential sources of error, as it is necessary that the S–S displacement of the glass slide should be reproducible on being replaced in the holder, and independent of any lateral motion of the slide while in the holder. In each of these respects the specimen holder was found to function extremely well, and no appreciable movement of the position of the glass surface or of a specimen surface has been observed on replacing them or on moving them laterally.

The S–S displacement of the slide in one set of results (curve 1, Fig. 2) could perhaps be ignored for most work. In the whole course of this work, however, the maximum S–S displacement observed was 0·04 mm, which corresponds to a maximum error of 0·02° in 2θ. Thus, while the shift in angle caused by the S–S displacement may in general be negligible, it can hardly be said to be so in any particular case unless it has actually been measured.

Conclusion

The S–S displacement of a specimen can be determined by this simple method to an accuracy of 0·006 mm. While the mechanical limitations with regard to the S–S displacement of the diffractometer used would appear to be somewhat severe, they can be overcome by taking certain simple precautions.

The measurement of the S–S displacement entails a certain amount of care, and this is quite useless unless similar care is taken with the preparation of the specimen. The specimen surface needs to be flat to at least 0·005 mm to make this procedure worth while; it is just worth while in the case of a specimen sieved through 300 mesh, where the average particle size is about 0·004 mm. Any variation in flatness of the specimen surface gives rise to variations in the S–S displacement and this leads to line broadening. The determination of the S–S displacement corresponding to the 'mean surface' of such a specimen would appear to be a very difficult task.

Acknowledgments

The author is greatly indebted to Dr. E. R. Pike for his suggestions and advice in the early part of this work, and to Professor A. J. C. Wilson and Dr. J. W. Hughes for their help and encouragement. The work has been carried out in connection with the author's employment by The Institute of Physics and The Physical Society on work on the A.S.T.M. X-ray Powder Data File.

References

PARRISH, W., 1960, *Acta Cryst.*, **13**, 847.

SWANSON, H. E., GILFRICH, N. T., and COOK, M. I., 1957, *Circ. Nat. Bur. Stand.*, 539, **7**, 1.

WILSON, A. J. C., 1950, *J. Sci. Instrum.*, **27**, 321.

Notes and comments

Notes for the Preparation of Papers for the *British Journal of Applied Physics*

The following instructions refer to the main points which have to be considered in the preparation of a paper for publication, and authors offering papers to the *British Journal of Applied Physics* for publication are asked to conform to these recommendations.

Manuscripts

Manuscripts should be typed in *double* spacing on paper not wider than 8 in. and not longer than 13 in. Only one side of the paper should be used, and a margin of 1–1½ in. should be left. As alterations in the text cannot be allowed once the paper is set up in type, authors should aim at absolute clarity of meaning and of typing, and should check the typescript carefully before submission. All manuscripts should be submitted in duplicate, and in addition to the fair copies, a set of small copies or prints of the diagrams (not larger than foolscap) must be attached to each MS. (This enables MSS. to be sent to two referees at the same time and so assists rapid publication.)

Abstract. An abstract is printed at the beginning of the paper immediately after the title, name(s) of author(s), and place of employment of author(s). Two extra copies of the 'title and abstract' page are required for the records.

Mathematics. It is not necessary to give detailed derivations of mathematical expressions and formulae in a published paper when the work is straightforward; it is quite sufficient to indicate the method of treatment and the final results.

References. In the text bibliographical references are made by giving the name of the author and the year of publication in brackets, e.g. (Jones 1942), and details are given in the last section, 'References', where the references are arranged in alphabetical order of authors' names and in date order for each author.

Drawings

Drawings should be in Indian ink on tracing cloth, tracing paper or white card, with lettering in soft or blue pencil; lettering of a size suitable for reduction will be inserted by our draughtsman. The drawings should in general be sufficiently large to allow of reduction in printing, and the lines should therefore be bold; the frame lines of graphs should be slightly finer than those of the plotted curves.

Full details are available in the Institute and Society's 'Notes for Authors', obtainable from the Editor and Deputy Secretary, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London S.W.7.

Papers for publication in the *British Journal of Applied Physics* should be sent to the Editor and Deputy Secretary at the same address.

Publication of Conference Reports

It is hoped in future to publish short reports of any Conferences which are likely to be of interest to readers of this journal.

The Editor would be very grateful if any members of the Institute and Society who will be attending a Conference would let her know if they would be willing to write a report on it. This should be up to 1000 words in length and should be a general survey of the Conference mentioning any particular points of interest.

British Journal of Applied Physics Editorial Board

Members of the Editorial Board for the British Journal of Applied Physics are as follows:

J. M. A. Lenihan, Ph.D., M.Sc., A.M.I.E.E., F.Inst.P. (*Chairman*).

Prof. R. W. Ditchburn, M.A., Ph.D., B.Sc., F.Inst.P. (*Ex-officio member*).

W. Hirst, Ph.D., B.Sc., F.Inst.P.

C. A. Hogarth, Ph.D., B.Sc., F.Inst.P.

J. W. Menter, M.A., Ph.D., F.Inst.P.

A. C. G. Menzies, M.A., D.Sc., F.Inst.P.

Prof. D. A. Wright, D.Sc., F.R.A.S., F.Inst.P.

Journal of Scientific Instruments

Contents of the August issue

PAPERS

A non-destructive method of gauging electrodeposited nickel layers. By P. M. Reynolds.
Measurement of magnetic field contours. By R. Freeman.
Automatic method of magnetic field calibration using proton resonance. By A. Horsfield, J. R. Morton, and D. G. Moss.
Ultra-violet absorption optical system with photoelectric recording for a Phywé ultracentrifuge. By J. B. T. Aten and A. Schouten.
Apertometer for microscope objectives. By C. J. D. Spencer and W. T. Welford.

RESEARCH NOTE

Suppressed zero d.c. voltmeter. By R. E. Aitchison.

LABORATORY AND WORKSHOP NOTES

Simple bridge for the direct measurement of temperature differences. By M. C. Godin.
Design of magnetoresistance elements. By D. H. Roberts.
On making thin 'Vinylite' V.Y.N.S. films. By D. W. Green.
A 'black-box' frequency stabilizer for an electron-spin-resonance spectrometer. By W. A. Gambling and T. H. Wilmshurst.
Simple approach to cryogenic wave-guide. By D. K. Milne.
Simple and cheap power supply for a Pointolite lamp. By R. E. Hayes and A. R. V. Roberts.
Rigid metal–glass vacuum systems. Strain relief and separation. By T. G. Cork.
Modifications to an M.S.2 mass spectrometer. By J. Cuthbert.
Transistorized water-level meter. By R. R. Vierhout.

CORRESPONDENCE

Aluminium bakeable vacuum seal. From L. Holland.

NOTES AND NEWS

New Instruments, materials and tools. Notes and comments.

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique).

All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.

REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611–2.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12*s.* 6*d.* each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.

CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

PRINTED IN GREAT BRITAIN BY UNWIN BROTHERS LIMITED, WOKING AND LONDON

The Birmingham conference on nuclear physics, April 1961

MS. received 5th June 1961

Abstract

The Institute of Physics and The Physical Society held a conference on nuclear physics in the buildings of the Departments of Physics and Mathematical Physics of the University of Birmingham on 17th, 18th and 19th April 1961. The subject-matter discussed covered a wide field of experimental and theoretical research.

By invitation of Professor P. B. Moon and Professor R. E. Peierls, The Institute of Physics and The Physical Society held a conference in the buildings of the Departments of Physics and Mathematical Physics of the University of Birmingham on 17th, 18th and 19th April 1961. Regular meetings of this type in nuclear physics are an important part of the Institute and Society's annual programme of conferences and they are usually arranged with the interests of younger research workers in the subject very much in mind. The organizers of the Birmingham meeting felt that it was right that the subject-matter discussed should cover a wide field of both experimental and theoretical research so that students with many different interests should hear something of each other's problems. Topics ranging from the Mössbauer effect to the C.E.R.N. proton synchrotron were consequently discussed before substantially the same audience.

For some years the Physical Society has customarily presented certain of its awards at Conferences. At the Birmingham meeting the President, Sir John Cockcroft, O.M., F.R.S., presented the Duddell Medal to J. B. Adams of C.E.R.N., Geneva, for his work on the design and construction of the proton synchrotron of that laboratory; he presented the C.V. Boys Prize to A. W. Merrison of Liverpool University for his work in high energy physics. In subsequent addresses they gave accounts of their work and its implications: these talks will be published (Merrison in full in *The Proceedings of the Physical Society*, and Adams, briefly in this journal), and will not be reported here.

At the beginning of the meeting it was reported that the 320 delegates (including 90 from the University of Birmingham) had submitted more contributions than could reasonably be accommodated, and some were therefore presented only as printed abstracts. It was also necessary to include certain parallel sessions. With these measures it was found possible to allocate sufficient time to each contributor for him to present a satisfactory account of his work and to allow time for discussion of the more controversial points. Each main session opened with a review paper and the succeeding short contributions were chosen, as far as possible, to illustrate points made in the review.

The first day of the conference was devoted mainly to the Mössbauer effect. W. Marshall (A.E.R.E.) reviewed the applications of recoilless emission of γ-radiation in solid state physics, referring in particular to the determination of the magnitude and direction of the internal magnetic fields in solid structures. The nuclear Zeeman effect in ^{57}Fe and ^{119}Sn has been used for this purpose. Present theory is unable to provide satisfactory estimates of the various contributions to these fields. It is helpful in this situation to supplement information obtained by the Mössbauer effect by nuclear magnetic resonance experiments on the effect of impurity atoms on the neighbouring nuclei of the host lattice. The Mössbauer lines also show a chemical shift (due to the electron density within the nucleus being dependent on the state of chemical combination) and a nuclear quadrupole shift. The speaker pointed out that in comparison with magnetic experiments, relatively little work had been done on the phonon spectrum of the lattice, but that it would be interesting to attempt observation of processes involving one or more phonons, in contrast with the zero-phonon Mössbauer transition. Short contributions on this topic included an account by A. J. F. Boyle (Manchester) of the nuclear Zeeman effect in ^{119}Sn and ^{161}Dy. H. E. Hall (Manchester) discussed the application of the Debye–Waller theory to resonant scattering and P. B. Moon (Birmingham) presented a theory of the interference between resonant and Rayleigh scattering of the Mössbauer line in ^{57}Fe. P. J. Black (Birmingham) described experimental evidence supporting this theory.

On the second day the first main subject was 'Low energy interactions' which was interpreted broadly to mean experiments and theories on nuclear structure. G. E. Brown (NORDITA) spoke on 'Vibrational states of nuclei'. He showed how, by a fusion of the ideas of the shell model and the collective model, the odd parity vibrational states could be considered to arise from interactions between a particle raised out of a closed shell and the resulting 'hole'. A large number of states are thus predicted, but the dipole states with isotopic spin 1, which are observed in giant resonance experiments with γ-radiation, are raised in energy by particle–hole repulsion. This accounts reasonably for the observed location of these states at about 20 MeV in the energy spectrum. The same sort of fusion of independent particle motion with collective motion was shown by E. B. Paul (A.E.R.E.) to give an excellent account of the low-lying energy levels of d-shell nuclei. The basic states are those described by Nilsson, in which a nucleon moves in a spatially non-uniform potential produced by a deformed nuclear core; rotational bands of levels are built upon these states.

A first group of shorter contributions to the topic of low energy interactions showed how nuclear states, including those of rotational bands, could be investigated with 150 MeV protons from a synchrocyclotron. The basic shell model states for both protons and neutrons can also be exhibited by studies of the (p, 2p) and (p, n) reactions at the same energy. In the case of the (p, 2p) reaction, an angular correlation measurement selects a particular momentum interval for the struck proton in the target nucleus and the distribution of these momenta can be correlated with the shell structure. The highly excited dipole states can be accurately examined through the (p, γ) reaction produced by protons from a tandem electrostatic generator; results for T(p, γ)^{4}He, ^{15}N(p, γ)O^{16} and ^{31}P(p, γ)^{32}S were reported.

Another group of papers dealt more with interactions than with nuclear states and showed how the angular and statistical accuracy of measurements of polarization in scattering processes had increased. Measurements of elastic and inelastic scattering of 10 MeV protons were reported and work on the small-angle scattering of neutrons of energy 18–120 MeV was also described. Several papers were devoted to the deuteron stripping reaction and served mainly to show how excellent a first approximation in many cases is the Butler theory of this process. Full correction for Coulomb and nuclear distortion effects does not necessarily produce better agreement with experiment and partial correction may easily make matters worse. The (d, n) stripping reaction usually employs the time-of-flight technique for definition of neutron energy and studies of the ^{40}Ca(d, n) and ^{10}B(d, n) reactions were reported. It is now becoming possible to make detailed predictions of cross sections under certain assumptions and it was shown how (*a*) low energy deuteron elastic scattering and (*b*) elastic and inelastic neutron scattering from ^{6}Li could be predicted from an optical model. From a survey of these contributions it is clear that accelerators such as the tandem Van de Graaff (~12 MeV), the proton linear accelerator (~50 MeV) and the synchrocyclotron (~150 MeV) each have much to offer the nuclear spectroscopist.

Towards the end of the second day a review paper on 'Collisions between nucleons at energies less than 1 GeV' was given by B. Rose (A.E.R.E.). He surveyed existing experiments and listed those still needed to obtain a complete description of the scattering matrix. The possibility of using polarized beams and polarized targets to simplify the determination of scattering parameters was also discussed. Experimental and theoretical effort in this field was illustrated by short papers from Birmingham on the isotopic spin dependence of nucleon–nucleon scattering amplitudes (600–1000 MeV) and from Liverpool on the neutron–proton scattering at about 180° in the centre-of-mass system. The high accuracy of the latter measurements was shown by Castillejo and Peierls (Birmingham) to permit a determination of the contribution made to the scattering by the charged pion pole. Good agreement was found with the estimate based on π-nucleon scattering. Reports were also given on the angular distribution of polarization effects in free neutron–proton scattering, on the spin–orbit forces between protons, and on optical model calculations of proton–proton and pion–proton elastic scattering. The higher energy region of interaction now made accessible by the operation of the C.E.R.N. proton synchrotron was illustrated by accounts of the production of hyperons in π^-–p collisions (C.E.R.N.) and of the analysis of π^-–p collisions at a momentum of 16 GeV/*c* (Birmingham–Imperial College–Oxford collaboration). The results of the latter experiment show qualitative agreement with the idea of an interaction with a 'peripheral' meson associated with the target proton. R. Mermod (C.E.R.N.) presented results on the total cross section for the interaction of protons, antiprotons and K mesons with protons, for the momentum range 4 to 15 GeV/*c*. The difference between the proton and antiproton cross sections is found to decrease over this interval, while the proton cross section itself stays constant.

On the last day of the Conference there were parallel sessions on instrumentation and machines, and on theoretical methods. In the former, G. H. Stafford (National Institute for Research in Nuclear Science) described the proton linear accelerator now operating at the Rutherford Laboratory at a mean energy of 30·1 MeV. A source of polarized protons has been installed and a current of 5×10^7 protons per sec with $P = 0{\cdot}31$ has been obtained at a full energy. The future of such accelerators was surveyed and it was pointed out that by the use of a superconducting tank liner, an enormous reduction in r.f. power losses could be achieved, and that a machine with a current of 1 mA of protons and 100% duty cycle could be envisaged. For 50 MeV the cost of refrigeration would be about £80 000. W. B. Powell (Birmingham) reported progress on the 40-inch radial ridge three-sector cyclotron and presented a design study for a high performance 60 MeV proton cyclotron which might compete favourably with the superconducting accelerator. High r.f. voltages on a three-sector dee and a large magnet gap are planned in this machine. Other machine developments, including ion storage in a synchrocyclotron (Harwell), and deuteron acceleration in a proton synchrotron (Birmingham), were reported.

The papers on smaller instruments included a description of the use of a 6 mm cube of silicon doped with phosphorus to form a diffused junction counter as a detector for 30 MeV protons. The surface barrier type of counter, furnished with a hydrogenous layer, has been used as neutron spectrometer for energies up to 7 MeV. High pressure ^{3}He ionization chambers and a liquid helium scintillation counter, and their applications, were also described. The use of boron-loaded phosphors in neutron flux measurements was discussed. Progress in the visual techniques was represented by a paper from Glasgow on a propane bubble chamber for the study of photoprocesses, in which a γ-ray beam traverses the chamber within a thin-walled tube containing the target, and a paper from Birmingham on the automatic measurement of bubble chamber tracks. The construction and use at Glasgow of triggered spark chambers containing He, Ne and Ar was also described.

The theoretical sessions dealt mainly with collective excitations of nuclei and with pairing forces, nuclear reactions and elementary particle processes. D. J. Thouless (Birmingham) showed how the time-dependent Hartree–Fock approximation has been used as the basis of a new approach to the problem of nuclear rotational states. The resulting expression for the nuclear moment of inertia is similar to the 'cranking model' result but contains correction terms which express the fact that the forces are really interactions between pairs of particles. These terms correspond to a selective summation of diagrams in perturbation theory. The method has been applied to ^{28}Si (Husain, Birmingham) using a non-singular saturating nuclear force and harmonic oscillator wave functions; the calculated value of the moment of inertia is close to the experimental result.

The relative importance of pairing forces and spin–orbit forces in the calculation of nuclear energy levels was examined (Umezawa, Strasbourg) and it was found that in nuclei for which *LS* coupling is good, the singlet pairing force dominates. In papers on nuclear reactions it was shown how the cross section for the ^{9}Be(^{3}He, t)^{9}B reaction could be reduced to that for n(^{3}He, t)p. There was also a description of the elimination of ambiguities in the Butler analysis of stripping reactions by the use of the distorted wave Born approximation. The polarization of protons elastically scattered by ^{7}Li near the 441 kev resonance was shown to determine the sign of the contributions of the two possible channel spins to the reaction cross section. Finally a discussion of inelastic processes in elementary particle interactions at high energies (Selleri, C.E.R.N.) showed that a remarkably good description could be obtained in terms of

the exchange of a single virtual particle. This 'peripheral' model has already been invoked successfully in the analysis of high energy experiments.

Delegates to the Conference were able to visit the laboratories of the Department of Physics informally throughout the meeting and slightly more formally on the second evening, when an 'At Home' with light refreshments was held. Some members were able to visit the Royal Shakespeare Theatre at Stratford on the same evening for a performance of *Much Ado About Nothing*.

University of Birmingham,
Edgbaston,
Birmingham 15.

W. E. BURCHAM
G. H. BURKHARDT
P. M. ROLPH

List of papers read at the Conference

1. The Mössbauer Effect in Dy^{161}
   A. J. F. Boyle, D. St. P. Bunbury and C. Edwards, University of Manchester.
2. The isomer shift in Sn^{119}
   A. J. F. Boyle, D. St. P. Bunbury and C. Edwards, University of Manchester.
3. The Debye–Waller factor for resonant γ-ray scattering
   H. E. Hall, University of Manchester.
4. Spectral distribution and coherence of gamma ray scattering near a resonance
   P. B. Moon, University of Birmingham.
5. Interference between Rayleigh and nuclear resonant scattering
   P. J. Black, D. Evans and D. O'Connor, University of Birmingham.
6. Inelastic scattering of 150 MeV protons from ^{16}O and ^{40}Ca
   D. J. Rowe, A. B. Clegg, G. L. Salmon, K. J. Foley and P. S. Fisher, University of Oxford.
7. Gamma-radiation from the 150 MeV proton bombardment of nuclei in the 2–1d shell
   K. J. Foley, A. B. Clegg, G. L. Salmon and D. J. Rowe, University of Oxford.
8. Binding energies of proton states in light nuclei
   H. G. Pugh, Atomic Energy Research Establishment and K. F. Riley, University of Cambridge.
9. Cross section of $N^{15}(p, \gamma)0^{16}$ in the giant resonance region
   G. C. Thomas and E. D. Earle, University of Oxford.
10. Neutrons emitted at 0^0 from nuclei bombarded by 143 MeV protons
    J. B. Scanlon, P. H. Bowen, G. C. Cox, G. B. Huxtable, A. Langsford and J. J. Thresher, Atomic Energy Research Establishment.
11. Polarization in $C^{12}(p, p)$ scattering in the energy range 5·5 to 10·5 MeV
    J. E. Evans, Atomic Energy Research Establishment.
12. Polarization of 10 MeV protons elastically scattered from copper
    A. B. Robbins, K. A. Grotowski and G. W. Greenlees, University of Birmingham.
13. Inelastic proton scattering asymmetry measurements at 8·9 MeV
    B. Hird, A. Budzanowski and R. W. Clifft, University of Liverpool.
14. Polarization in neutron–deuteron elastic scattering up to 1·0 MeV
    R. E. White and A. T. G. Ferguson, Atomic Energy Research Establishment.
15. Small angle scattering of neutrons by uranium at energies between 18 MeV and 120 MeV
    A. Langsford, P. H. Bowen, G. C. Cox, G. B. Huxtable, J. P. Scanlon and J. J. Thresher, Atomic Energy Research Establishment.
16. Deuteron stripping reactions of low Q values
    D. W. Mingay, Atomic Energy Research Establishment.
17. A study of the $B^{11}(d, n\gamma)C^{12}$ reaction
    J. B. Garg, N. H. Gale and J. M. Calvert, University of Manchester.
18. The angular distributions of alpha-particle groups from the reaction of $^{24}Mg(^3He, \alpha)^{23}Mg$
    S. Swierszczewski, G. Parry and H. D. Scott, University of Liverpool.
19. Calculations for the (p, d) reaction
    P. Gould, University of Birmingham.
20. A study of the reaction $^{40}Ca(d, n)^{41}Sc$
    B. E. F. Macefield, J. H. Towle and W. B. Gilboy, Atomic Weapons Research Establishment.
21. Energy levels of C^{11}
    A. N. James, University of Cambridge, A. T. G. Ferguson, Atomic Energy Research Establishment, and C. M. P. Johnson, University of Oxford.
22. Shell model of B^{11}
    J. M. Soper, Atomic Energy Research Establishment.
23. Optical model analysis of the elastic scattering of low energy deuterons
    R. N. Maddison, University of Oxford.
24. The inverse photo-disintegration of He^4 and S^{32}
    D. S. Gemmell and G. A. Jones, Atomic Energy Research Establishment.
25. Neutron scattering from Li^6
    D. F. Jackson, Battersea College of Technology.
26. The $^{12}C(p, p')^{12}C$ and $^{12}C(p, \gamma)$ reactions from 10–30 MeV
    D. F. Measday, P. S. Fisher and A. B. Clegg, University of Oxford.
27. Isotopic spin dependence of nucleon–nucleon scattering amplitudes between 600 MeV and 1000 MeV
    R. Rubinstein, G. Martelli, H. B. van der Raay, K. R. Chapman, J. D. Dowell, W. R. Frisken, B. Musgrave and D. H. Reading, University of Birmingham.
28. Investigation of the production of ground state and excited hyperons in π^-–p collisions
    J. D. Dowell, B. Leontic, A. Lundby, R. Meunier, G. Petmezas, J. P. Stroot and M. Szeptycka, C.E.R.N.
29. π^-–p interactions at 16 GeV/c
    B. Tallini, University of Birmingham.
30. Polarization effects in neutron–proton scattering at 150 MeV
    C. Whitehead, Atomic Energy Research Establishment.
31. Proton–proton spin-orbit interaction and a possible vector meson pole term
    D. W. L. Sprung, University of Birmingham.
32. The measurement of the neutron distribution near the backward direction in n–p scattering at 353 MeV
    W. H. Range, A. Ashmore, R. T. Taylor and B. M. Townes, University of Liverpool.
33. Extrapolation of n–p scattering to the one pion pole
    L. Castillejo and R. F. Peierls, University of Birmingham.
34. Nucleon–antinucleon annihilation to two mesons
    R. J. N. Phillips, Atomic Energy Research Establishment.
35. Optical model calculations of p–p and π^-–p elastic scattering
    D. V. Bugg, University of Cambridge.

36. An accurate determination of the proton linear acceleration beam energy by a time of flight method
D. J. Warner and C. J. Batty, N.I.R.N.S.
37. An automatic track following device
M. Jobes and G. A. Doran, University of Birmingham.
38. Preliminary experiments with a solid state ionization chamber
C. J. Batty and R. J. Griffiths, N.I.R.N.S.
P. E. Gibbons and D. C. Northrop, Services Electronics Research Laboratory.
39. A semiconductor proton-recoil neutron spectrometer
G. Dearnaley and A. T. G. Ferguson, Atomic Energy Research Establishment.
40. Triggered spark chambers
J. M. Paterson, J. G. Rutherglen and B. H. Patrick, University of Glasgow.
41. A bubble chamber for γ-ray experiments
J. M. Scarr, I. S. Hughes and R. Jennings, University of Glasgow.
42. The application of a helium-3 ionization chamber to the study of neutron spectra from (p, n) reactions
J. M. Freeman and D. West, Atomic Energy Research Establishment.
43. A liquid helium scintillation counter
A. M. Segar, H. L. Anderson, E. P. Hincks, C. Rey and C. S. Johnson, University of Chicago.
44. The characteristics of a small neutron detector
J. B. C. Brown and S. A. Scott, University of Birmingham.
45. The radial ridge cyclotron with injection of polarized deuterons
W. B. Powell, A. J. Cox, T. W. Eaton, E. S. Finlay, D. E. Kidd, B. L. Reece and P. J. Waterton, University of Birmingham.
46. The Birmingham synchrotron: a new radio-frequency system and the acceleration of deuterons
P. D. Whitaker and H. R. Shaylor, University of Birmingham.
47. On increasing the beam intensity of existing synchrocyclotrons
F. M. Russell, N.I.R.N.S.
48. Proposal for a cyclotron to give 60 MeV protons at high intensities
W. B. Powell, P. J. Waterton, University of Birmingham.
49. On the nuclear moments of inertia
D. Husain, University of Birmingham.
50. Translation modes of self-bound systems
D. E. McCumber, University of Copenhagen and C. E. Brown, Nordita.
51. Competition between pairing forces and the spin orbit force in nuclei
M. Umezawa, Strasbourg.
52. A study of the charge exchange $Be^9(He^3, t)B^9$
J. E. Young, University of Copenhagen and P. R. Stein, Los Alamos.
53. Removal of l-value assignment ambiguities in the analysis of (d, p) stripping angular distributions
H. D. Scott, University of Liverpool.
54. The inelastic scattering of alpha-particles by the nuclear surface interaction
G. Mandel, University of Oxford.
56. Polarization of protons elastically scattered by Li^7 near the 441 kev resonance
J. R. Rook, University of Manchester.
56. Peripheral model for inelastic processes
E. Ferrari and F. Selleri, C.E.R.N.

Supplementary list of papers submitted (but not read because of lack of time)

1. Fine structure of μ-mesic x-rays
A. M. Segar, H. L. Anderson, E. P. Hincks, C. S. Johnson, University of Chicago.
2. The neutron cross section of thorium between 300 ev and 30 kev
C. A. Uttley and R. H. Jones, Atomic Energy Research Establishment.
3. The D(d, p)T cross section from 4 to 15 kev
C. C. Goodyear, University of Oxford.
4. Evidence for a stripping mechanism in lithium induced nuclear reactions
G. C. Morrison, University of Chicago.
5. Photonuclear cross section in molybdenum
K. H. Lokan, Atomic Energy Research Establishment.
6. The levels of B^{11} from the $Li^7(\alpha,\gamma)B^{11}$ reaction
G. A. Stephens, L. L. Green, J. C. Willmott, University of Liverpool.
7. Decay modes of the first six excited states of Mg^{24}
J. A. Cookson, A. V. Cohen, Atomic Weapons Research Establishment.
8. Resonances in the $K^{39}(p, \gamma)Ca^{40}$ reaction from 1·0 to 2·1 MeV
R. A. Pope, D. V. Freck and W. W. Evans, A.E.I., Aldermaston.
9. Two-quantum decay in Zr^{90}
R. J. Ellison and B. Dickinson, University of Manchester.
10. Distribution of partial radiation widths
J. R. Bird, Atomic Energy Research Establishment.
11. Inelastic scattering of 150 MeV protons from ^7Li and ^{11}B
A. B. Clegg, G. L. Salmon, D. Newton, K. J. Foley, P. S. Fisher and D. J. Rowe, University of Oxford.
12. The $^{12}C(p, 2p)^{11}B$, $^{12}C(p, pn)^{11}C$ and $^{12}C(n, np)^{11}B$ reactions induced by nucleons of 120–150 MeV
G. L. Salmon, S. M. Austin, D. J. Rowe, A. B. Clegg and D. Newton, University of Oxford.
13. An analysis of the $C^{12}(p, 2p)B^{11}$ reaction at 153 MeV
J. Nuttall and K. F. Riley, University of Cambridge.
14. Inelastic scattering of fast protons
E. A. Sanderson, University of Birmingham.
15. Operation of the 9-inch bubble chamber with deuterium
W. P. Dodd, G. A. Doran, J. B. Kinson, L. Riddiford and B. Tallini, University of Birmingham.
16. Total cross sections for p, $\bar{p}$ and $K^{\pm}$ on hydrogen between 4 and 15 GeV/c
G. von Dardel, R. Mermod, P. A. Piroué, M. Vivargent, G. Weber and K. Winter, C.E.R.N.
17. Emission of Li^8 from highly excited nuclei
S. A. Durrani, University of Oxford.
18. The effect of hard core and three-body forces in the λ-nucleon interactions
Mrs. L. F. Abou-Hadid, University College, London.

Summarized proceedings of a conference on X-ray analysis —Glasgow, April 1961

Abstract

The Spring Conference of the X-ray Analysis Group of The Institute of Physics and The Physical Society was held in the Chemistry Department of Glasgow University, on 6th and 7th April 1961. *Dr. C. W. Bunn, Chairman of the Group, presided at the first and third sessions, whilst Professor J. M. Robertson, F.R.S., was in the chair during the second session. The Evening Discourse on 6th April was given by Professor R. V. Jones (Aberdeen University), who discussed some aspects of the direction of the national scientific effort in connection with the war of* 1939–45 *under the title 'Philosophers and Kings'.*

THE subject proposed for the Conference was 'The Structures of Organic Compounds', but this phrase had been interpreted broadly, and the seventeen papers presented covered a range of topics, not always closely related. This summarizing report therefore covers the proceedings in an arrangement suggested by subject matter, rather than dealing with each paper and its subsequent discussion in chronological order.

In principle we should be able to predict all the details of a crystal structure, and particularly the manner of molecular packing, if we had an adequate knowledge of the relevant interatomic and intermolecular forces. In fact our knowledge is far from adequate, and Dr. S. C. Nyburg (University College of North Staffordshire), in a paper entitled 'Structure and lattice energy of molecular crystals', sketched a simplified approach to this difficult problem. He stressed the desirability of considering the structure at low temperatures, since polymorphic changes may supervene, with the result that the room-temperature structure may differ from that predicted for minimum free energy. Intermolecular forces were assessed by a Lennard-Jones spherical-atom formula for each atom; from this, for a number of diatomic (or quasi-diatomic) molecules such as H_2, Cl_2, Br_2, I_2, N_2, O_2, CO_2 and C_2H_2, he attempted to determine the mode of packing that would minimize the free energy. He was thus able to explain certain features of the known crystal structures, notably the mutual orientation of the molecular axes. There were indications that the molecular quadrupole moment may sometimes exert a decisive influence on the structure. In discussion Dr. D. R. Holmes mentioned some calculations by Mr. I. Barrie, who had been able to account for the molecular tilt in the orthorhombic modification of long-chain hydrocarbons.

'Imferron' is a medicinal material made by producing ferric hydroxide in presence of dextran. Because of the protective action of the latter, the effective solubility of the hydroxide becomes very high. In a paper, 'X-ray diffraction by an iron-dextran complex', Dr. J. Iball (with Mr. C. H. Morgan, Queen's College, Dundee) described studies of this material both as fibres drawn from the highly viscous liquid and as 'powder' from the dried fluid. The diffraction pattern showed several rather well-defined lines, the most prominent corresponding to a spacing of 2·56 Å. These lines agreed with some of those observed in the diagram for colloidal ferric hydroxide and it seemed likely that the complex consists of micelles embodying well-orientated crystallites with dimensions of the order of 100 Å.

Dr. D. Lawton (Glasgow University) reported 'Studies of some inclusion compounds' of tri-*o*-thymotide (TOT). Those described hitherto have belonged to two general types: (*a*) cavity structures, in which the TOT crystallizes trigonally so as to leave cages which entrap molecules up to 9·5 Å in length, and (*b*) channel structures, which may be trigonal or hexagonal, capable of accommodating longer molecules. Dr. Lawton described a number of compounds which do not fit into the above classification, notably one formed with trimethylene dibromide which crystallizes in the orthorhombic system and is yet another example of spontaneous resolution, and one formed with carbon tetrabromide which differs from that enclosing the slightly smaller molecule of carbon tetrachloride. The included molecules may not always conform to the space group adopted by the TOT structure. In the study and analysis of these adducts, an accurate determination of the density is valuable, and a gradient-tube is being used for this purpose.

Large molecules of proteins and viruses may comprise a number of sub-units which are identical or nearly so, but differently orientated. The Patterson function of the parent crystal will then include a superposition of the Patterson functions of the sub-units in different orientations. Dr. M. G. Rossmann (with Dr. D. M. Blow, MRC Unit for Molecular Biology, Cambridge), in a paper on 'The relative orientation of sub-units in protein molecules', examined some of the consequences of this principle, and showed that it might lead to a solution of the phase problem, provided the Patterson function could be satisfactorily resolved into its components. As a first step in this scheme, Dr. Rossmann had written an EDSAC programme which took the observed three-dimensional Patterson function and superposed it upon itself in relative orientations that could be systematically varied, and then evaluated a function designed to measure the goodness of fit at each orientation. Positions at which this function is a maximum should correspond to the relative orientations of the sub-units in real space. This programme had been tested on the experimental results for haemoglobin, and it satisfactorily verified the known relative orientation of the two halves of the protein molecule.

Dr. B. Stevens (Sheffield University) read a short paper on 'Fluorescent spectra and crystal structure of aromatic hydrocarbons'. In solution the spectrum of pyrene changes character as the concentration rises: the ordinary structured band diminishes, whilst a structureless band at a different frequency becomes dominant, this latter band being attributable to a dimeric molecule formed by union of a normal and an excited monomeric molecule. This same band appears in the fluorescent spectrum of solid pyrene, which is known to have an unusual type of crystal structure (Fig. 1, B) with pairs of parallel molecules instead of the one (A) adopted in most aromatic hydrocarbons (e.g. anthracene), which give the normal type of spectrum. Perylene also crystallizes with a type B structure, and gives the anomalous spectrum, and this correlation appears to be general. In discussion with

Dr. M. R. Truter, Dr. Stevens made the point that this dimerization depends on one molecule's being excited, so that it could not be detected by ordinary methods for measuring molecular weight.

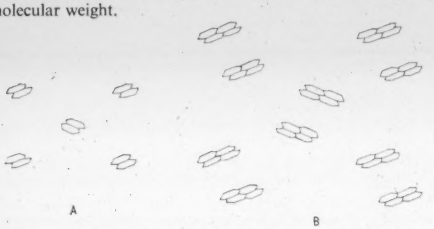


Fig. 1. Two types of packing of planar aromatic molecules in crystals.

Analysis of the crystal structure of a complex organic crystal proceeds by reiteration of a sequence of steps: Fourier synthesis, scanning of the electron density distribution to locate—and ultimately to identify—significant peaks, allocation of new atomic coordinates and vibrational parameters, calculation of structure factors with phase angles, and use of the latter in a new Fourier synthesis. Ordinarily these steps are taken separately, results being printed out and considered at each stage. In a paper entitled 'Automatic Fourier analysis of a complex organic crystal', Dr. J. S. Rollett (with Mr. L. I. Hodgson, Oxford University) outlined a plan for carrying out the whole cycle automatically without necessarily inspecting the intermediate results. The large amount of storage space needed in the computer imposes severe restrictions on complete automatization. But this 'progress report' described successful analyses of histidine hydrochloride of aneurin phosphate, and—more significantly—of the partially unknown structure of a coordination compound with gross formula, $PtIAs_4BC_{78}$. In this last the positions of the heavier atoms only were known beforehand, yet the programme found and correctly identified the lighter atoms (other than H) and indicated the more specific formula,

$$[PtI\{As(C_6H_4AsPh_2)_3\}]^+ [BPh_4]^- .$$

Discussion, with Professor J. M. Robertson and Dr. O. S. Mills, concentrated on the problems which arise when initially incorrect phase-angles or incomplete and faulty intensity data cause the atomic peaks to be of poor shape, and hence ill-adapted for machine study. It appears that these difficulties are not insuperable, though the electron density count within a peak may, in the early stages of analysis, deviate widely from its theoretical value.

The remaining papers dealt with specific structure analyses, and they will be summarized in order of increasing molecular complexity. Dr. R. W. H. Small (with Mr. B. Beagley and Mr. D. O. Hughes, Birmingham University) described an accurate analysis of the 'Structure of 2-monofluoroacetamide', $CH_2F.CONH_2$. Three-dimensional (proportional-counter) intensity data were the basis for a least-squares refinement, and the final bond lengths, after correction for the effects of torsional oscillation of the molecule, had a standard deviation of about 0·005 Å. (In this hydrogen-bonded structure the centre of libration was displaced by about 0·25 Å from the molecular centre of mass.) C–C = 1·533, C–F = 1·401, C–N = 1·319, C–O = 1·256 Å. The molecule, apart from the H-atoms which were located, is very nearly planar. Brief mention was also made of an accurate study of ammonium oxamate $NH_4^+ (O_2C.CONH_2)^-$ in which the oxamate ion is nearly planar, and the ammonium ion appears to be 'anchored' by hydrogen bonds so that it is static instead of having the usual 'free rotation'. C–C = 1·556 ± 0·003 Å.

Mr. G. Ferguson (with Dr. G. A. Sim, Glasgow University) described 'An x-ray study of steric effects in some substituted benzoic acids'. Benzoic acid itself has the carboxyl group group very nearly coplanar with the benzene ring; but substituents in the *ortho*-positions cause overcrowding, which is minimized chiefly by the twisting of the carboxyl group out of the ring-plane, and to a lesser extent by the deviation both of this group and of the substituent from thejr ideal positions on two-fold axes of the benzene ring. In the *o*-chloro-acid the twist is 14°, whilst in the bromo-acid (I of Fig. 2) it rises to 18°. The twist is still greater, 23°, in the nitro-derivative (II), and it appears that the effect of the nitro-group is relayed, via the intervening H-atom, to the carboxyl-group. In the molecule III, it is the nitro-group which is twisted (through more than 40°) more than the carboxyl (9°). In reply to a question by Dr. Small, Mr. Ferguson said that the exo-cyclic C–C bond in *o*-chlorobenzoic acid came out at 1·521 ± 0·009 Å.

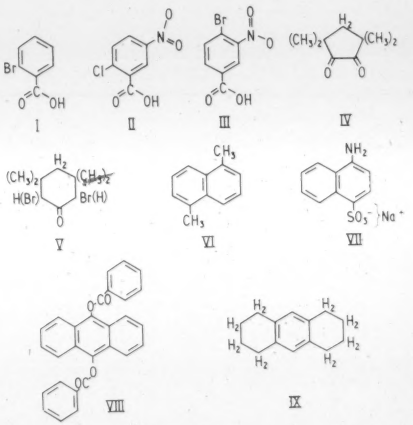


Fig. 2.

Miss L. Goaman (with Dr. D. F. Grant, University College, Cardiff) reported on 'The crystal structures of 2,2,5,5–tetramethyl–3,4–diketocyclopentane and 2,6–dibromo–3,3,5,5–tetramethyl–cyclohexanone', compounds whose molecules might show effects due to overcrowding of the substituents. The former (IV) molecule adopts strict *mm*-symmetry, so that only one parameter is needed to define the position of the molecule in the tetragonal cell. However, the crystals were difficult to handle, and the analysis was not highly refined. The latter compound (V) had been taken to an advanced state of refinement, though the analysis was not yet complete. The six-membered ring adopts a distorted 'chair' conformation.

Dr. J. Beintema (Shell Laboratories, Amsterdam) described a two-dimensional analysis of 'The crystal structure of 1,5–dimethylnaphthalene', a molecule (VI) which might show slight overcrowding. In fact no evidence of this was obtained. The two (crystallographically non-equivalent) C–CH_3 distances were 1·49 and 1·52 Å. The hydrogen atoms were

revealed in difference syntheses, whence it was concluded that the methyl groups do not here have 'free rotation'.

Dr. C. J. Brown (with Dr. D. E. C. Corbridge, I.C.I., Dyestuffs Division, Manchester) read a paper on 'The crystal structure of sodium naphthionate and the crystallography of related compounds'. After a survey of a large number of naphthionic acid salts, the tetrahydrated sodium naphthionate (VII) was selected for detailed analysis by three-dimensional methods, and its structure has been highly refined. S–C = 1·766, S–O(mean) = 1·45, and C–N = 1·41 Å, whilst within the naphthalenic residue the bond lengths were remarkably similar to those in naphthalene itself. Each Na ion was surrounded by six O atoms at the corners of fairly regular ocathedra, these having one edge in common. Though not isomorphous, calcium naphthionate has a similar structure, despite the fact that two Na ions have been replaced by a single Ca ion. Possibly this ion occupies one octahedron, leaving a vacant hole in the other.

In 'The use of three-dimensional and optical transforms in solving the structure of 9,10–anthrahydroquinone dibenzoate' by Dr. K. J. H. Mackay (with Dr. J. Iball, Queen's College, Dundee), the interest lay mainly in the method used for solving the crystal structure. Since the substance crystallizes with only one molecule in the unit cell, the molecular Fourier transform can be recognized in a relatively simple way in reciprocal space; and, since the molecule consists essentially of two parts (anthracene residue and benzoate groups), each planar but inclined to one another, its transform consists of correspondingly inclined 'columns' in reciprocal space. As was demonstrated with a Perspex model, this pattern can be identified in the weighted three-dimensional reciprocal lattice. The structure thus solved (VIII) was refined by least-squares methods.

In relation to this paper and several others, there was some discussion of the weighting system to be used in least-squares refinement. Dr. D. W. J. Cruickshanks (Leeds University) stressed the fact that different weighting-systems could lead to significantly different bond lengths and to very different assessments of their standard deviations. It was important to check the system, and two common ways of doing this were not equivalent: (*a*) consideration of the consistency between independent experimental determinations of intensity, and (*b*) study of the trends of $w\Delta^2$ with intensity and with $\sin\theta$.

'The crystal structure of 1,2,3,4,5,6,7,8-octahydroanthracene (IX)' was first studied by Dr. H. Mendel (with Miss L. H. Toneman and Mrs. J. P. M. A. van Asselt, Shell Laboratories, Amsterdam) with Cu-radiation at room temperature. Though the two-dimensional analysis was refined to an impressive R value of 9%, some of the bond lengths found were implausible. Therefore the intensities were redetermined, using Mo-radiation at −100° C, and the more complete data led to a much better analysis, though the R values were higher (16% for 200 *h*0*l*, and 19% for 150 *hk*0 reflections). In the reduced rings, which adopt a semi-'chair' conformation, the C–C bonds were longer (1·50–1·52 Å) than in the benzenoid ring (~1·42 Å). The H atoms were located in difference projections.

The last four papers to be summarized dealt with more complex molecules of biological importance. Dr. A. Vaciago read a paper 'On the determination and refinement of the crystal structures of the related antibiotic compounds, sodium cephalosporin C, 6-amino-penicillanic acid, and potassium benzylpenicillin', covering work done in collaboration with Dr. R. D. Diamand, Dr. E. N. Maslen, Dr. J. S. Rollett and Professor D. Hodgkin (Oxford University). The main effort had been devoted to solving the structure of the first-named (see Fig. 3). The phase problem had to be approached by way of the S atom, though this atom is inadequately heavy for a structure of this size. Intense sharpening had to be applied to the Patterson function before S . . . S vectors could be identified with certainty. Once this identification was secure, analysis proceeded by a combination of several different strategies, leading to the electron density synthesis shown in Fig. 3, with $R = 26{\cdot}4\%$ for some 1400 reflections.

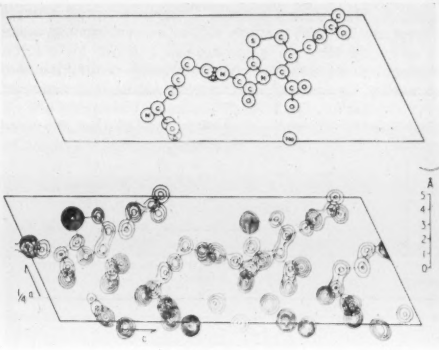


Fig. 3. Structure of cephalosporin C (sodium salt), including its absolute configuration, and an electron density map, made from sections through the three-dimensional function near the atomic centres.

Since the asymmetric unit comprises two similarly, but not identically, orientated molecules, a preliminary analysis on the basis of a pseudo-cell with $c' = c/2$ gave a fairly good 'average' structure.

The experience gained thus was applied to the study of the simpler structure of the aminopenicillanic acid, where only 14 atoms (not H) had to be located. This had been refined to $R = 10\%$ for about 1000 reflections. Dr. Vaciago ended with a brief account of the use of fuller experimental data and modern methods of computation in a refinement of the already known structure of potassium benzylpenicillin. This had now been taken to a stage where the standard deviation of some bond lengths was as low as 0·02 Å.

At Glasgow during the past twenty-eight months, structures have been ascertained for about a dozen rather complex molecules that are of key importance in natural-product chemistry. In a paper entitled 'A survey of recent x-ray studies of natural products by the Glasgow crystallography group', Professor J. M. Robertson outlined the strategic principles used, and discussed some of the implications of this development as it affects the classical problem of structure-determination in organic chemistry. He stressed the value of the *chemical* approach: the problem was seen not as the structure of a particular crystal, but rather as that of a series of chemically inter-related molecules. It was the chemist's task to determine the structural relationship between a number of different substances, and the crystallographer's to determine the exact structure of whichever of them proved most suitable for analysis. 'Better three months of chemical work than twelve of unrewarded x-ray effort.' He illustrated this concept by special reference to limonin: from several possible

derivatives, the iodoacetate of *epi*-limonol was finally selected for attack, despite the adverse circumstance that the asymmetric unit contained two molecules, so that 76 atoms had to be located instead of 38. This had the initial advantage of avoiding the phase ambiguity which otherwise occurs in $P2_1$, and the subsequent one that the finding of chemically identical structures for two crystallographically independent molecules afford very convincing evidence of their validity. On the other hand, the analysis of the simpler molecular structure of isoclovene hydrochloride had to be made despite the presence of the phase ambiguity.

Dr. C. H. Carlisle (Birkbeck College) instanced α-amyrin phosphate, in which not only did the asymmetric unit consist of two molecules, but whose analysis had also to be forced through against a virtual phase ambiguity of the kind mentioned by Professor Robertson. In subsequent discussion with Dr. Rollett, Dr. Mills (Manchester University) and Dr. Cruickshank, Professor Robertson agreed that the practicability of the strategy described depended on the prior development of fast and reliable electronic computers and on the availability of efficient crystallographic programmes for use on the computers. The establishment of a large programme is a most intricate and arduous task, and one which is all too often taken for granted, once the programme is working satisfactorily.

The chemical approach was further illustrated in papers by Dr. T. A. Hamor and Dr. G. A. Sim (both of Glasgow University) on 'Recent work on certain alkaloids', and on 'Recent work on some terpenoid derivatives', respectively. The former dealt with echitamine bromide and macusine-A iodide; the latter with bromogeigerin, clerodin, and certain chloronitrosocamphanes. The stereochemistry of these last compounds is of special interest in relation to the Cotton effect.

Chemistry Department,
The University,
Glasgow, W.2.

J. C. SPEAKMAN

Thermionic generation of electricity

by M. A. CAYLESS, B.Sc., A.R.C.S., A.Inst.P., Research Laboratory, Associated Electrical Industries (Rugby) Ltd., Rugby, Warwickshire

MS. received 16*th May* 1961

Abstract

A thermionic generator of electricity is essentially a diode valve in which electrons emitted from a hot cathode flow to a cooler anode, producing an electric current. From being scientific curiosities, such devices have become the subject of intense research activity in the last three years, and it is now clear that they have considerable possibilities as useful generators in a number of fields.

This article reviews this recent work and makes an assessment of the present position and future trends. Although engineering design and applications are considered, the emphasis is on the physical processes associated with these devices, and the progress which has been made into understanding them.

1. Introduction

THE production of electricity directly from energy sources without the use of moving machinery has become the subject of intense research activity in the last three or four years. Devices under investigation include thermoelectric and thermionic generators, photovoltaic cells, magnetohydrodynamic generators and fuel cells. Only the first two of these generate directly from heat, and only thermionic generators will be considered in this article (thermoelectric generators have been the subject of several comprehensive reviews, e.g. Jaumot 1958, Goldsmid 1960, Egli 1961). For a summary of other methods, see Spring (1961).

A thermionic generator is essentially a diode valve in which electrons from a hot cathode flow to a cooler anode, producing an electric current which can be used in an external circuit. Usually a gas or vapour is introduced to provide ions to neutralize the space-charge of the electrons, and devices of this sort have been variously termed thermionic converters, thermoelectron engines, plasma diodes, plasma thermocouples, etc.: the term 'thermionic generator' will be used here, together with 'caesium diode' for a particular type (section 5).

That electric currents can be generated in this way has been known since the last century: it is occasionally called the 'Edison effect'. Schlichter (1915) is quoted by Wilson (1960) as having proposed a generator of this form, and Champeix (1951) gave a quite detailed analysis which is essentially correct, although he overlooked several possibilities and concluded that it was impractical. In the same year Medicus and Wehner (1951) made a discharge tube which generated in this way, and the same effect was also observed by Malter, Johnson and Webster (1951). In 1954 Hurst (1958) applied for a patent covering generators of this type. Further early work in the U.S.S.R. is quoted by Dobretsov (1960) and Morgulis and Naumovets (1960).

It was not, however, until 1957 that accounts of systematic investigations into the practical realization of useful generators began to appear. Since the early paper by Moss (1957), several dozen original papers and several times this number of surveys, digests, etc., have been published, almost exclusively from the U.S.A. and U.S.S.R. In addition over half the book of collected papers on *Direct Conversion* by Kaye and Welsh (1960) is devoted to thermionic generation. This article aims at surveying this considerable output and reviewing the present position, with emphasis on physical principles rather than engineering and applications, which are amply covered elsewhere (e.g. Harvard 1959, *Electronics* 1960 b, Yaffee 1959, Grattridge 1960).

2. Principles

2.1. *Generation of power*

The anode current characteristic of an ordinary vacuum diode has the form shown in Fig. 1(*a*). The current in the

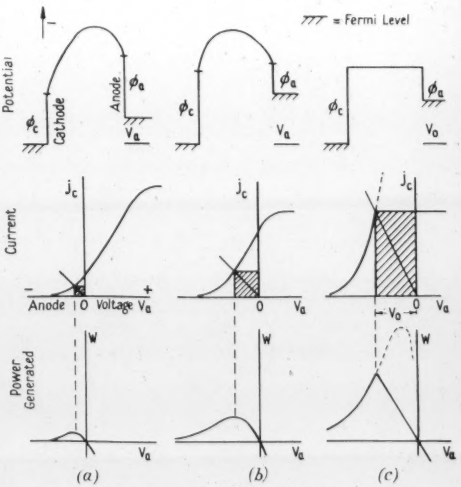


Fig. 1. Form of potential distribution and characteristics (not to scale): (*a*) ordinary vacuum diode, (*b*) $\phi_a < \phi_c$, (*c*) space charge completely neutralized (broken line: effect of reducing ϕ_c). In each case the potential diagram corresponds to the condition giving maximum power, indicated by the load lines and shaded areas shown on the characteristics.

retarding region represents a generation of electricity. With a particular external resistance the power generated is proportional to the shaded area enclosing the load line: this is plotted below and is a maximum when the load is 'matched'.

In an ordinary valve of this kind the anode work function ϕ_a is greater than the cathode work function ϕ_c, and there is an electron space charge between the electrodes, giving a potential diagram like that shown above the characteristic. The current generated is produced by those electrons from the hot cathode which are energetic enough to surmount both the work function ϕ_c and the space-charge potential barrier.

The effectiveness of the generator is obviously improved if $\phi_a < \phi_c$, as in Fig. 1(*b*). This shifts the characteristic to the left, and the power generated is considerably increased. If, in addition, the space-charge barrier is removed, still more power results, as shown in Fig. 1(*c*).

An effective thermionic generator is thus a diode valve in which:

(*a*) the anode work function is reduced, relative to the cathode work function;
(*b*) the space charge between the electrodes is reduced or eliminated.

The construction must be such that heat can be supplied to the cathode and removed from the anode. The ordinary glass bulb construction is not suitable (except perhaps with focused radiant energy heat sources) and some such construction as shown schematically in Fig. 2 is the simplest conceivable.

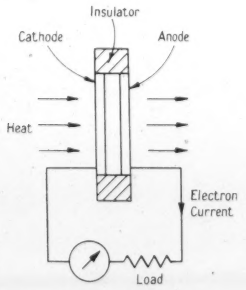


Fig. 2. Schematic thermionic generator.

It is readily shown (e.g. Feaster 1958) that, with complete space-charge neutralization, maximum power is produced when the load is matched so that the generated voltage

$$V_a = \phi_c - \phi_a, \tag{1}$$

provided

$$\phi_c - \phi_a > \frac{kT_c}{e}\left(= \frac{T_c}{11\,600}\ \text{volt}\right), \tag{2}$$

where T_c is the cathode temperature. This is the situation represented by the potential diagram and load line shown in Fig. 1(*c*).

Further power could be obtained by reducing ϕ_c until (2) became an equality (keeping $\phi_a + V_0 = \phi_c$, and T_c constant), as indicated by the broken line in Fig. 1(*c*), but only by much reducing the impedance and efficiency (see next section). If ϕ_c is reduced, whilst maintaining $V_a + \phi_a$ and T_c both constant, the power is unchanged: this is of little practical use, however, since the lower the work function the more difficult it is to maintain a cathode stable at a given temperature.

Future discussion is therefore confined to generators working in the optimum condition shown in Fig. 1(*c*), and complying with (1) and (2).

2.2. *Efficiency*

The device is, of course, a heat engine, the source of power being the temperature difference $T_c - T_a$ between the cathode and anode. Thermodynamically, therefore, the efficiency η, which cannot in any case exceed $1 - \phi_a/\phi_c$, is always less than $1 - T_a/T_c$. Richardson's law shows that if $T_a/T_c = \phi_a/\phi_c$, back emission from the anode is equal to the emission from the cathode, and the device acts as a reversible engine in this condition. To extract a finite amount of power, T_a is reduced somewhat, so that in operation

$$\eta \leqslant 1 - \frac{\phi_a}{\phi_c} < 1 - \frac{T_a}{T_c}. \tag{3}$$

It is convenient to define

$$\eta_c = 1 - \frac{\phi_a}{\phi_c} \tag{4}$$

as the 'Carnot efficiency', the highest conceivable efficiency for any given cathode–anode combination. The realizable efficiency falls below this, mainly because of heat transfer by radiation from the cathode to the anode. Denoting this by P_r (per unit area), and the current density by j_c, the maximum realizable efficiency is

$$\eta_m = \frac{j_c(\phi_c - \phi_a)}{j_c\phi_c + P_r} = \eta_c \Big/ \left(1 + \frac{P_r}{j_c\phi_c}\right) \tag{5}$$

or

$$\eta_m = \eta_c \Big/ \left(1 + \frac{1}{M}\right), \tag{6}$$

where

$$M = j_c\phi_c/P_r \tag{7}$$

is a convenient figure of merit for a cathode, a high value indicating a high realizable efficiency.

Thus for high efficiency it is desirable that ϕ_a is substantially less than ϕ_c, the current density is high and that the radiation from the cathode is small. At first sight a high value of ϕ_c also appears advantageous, but the current density j_c depends exponentially on ϕ_c, and there is an optimum value of ϕ_c for maximum efficiency under any given conditions.

P_r is difficult to estimate with precision, but it never exceeds the power loss determined by Stefan's law, which thus gives a minimum value for M of

$$M_{min} = \frac{j_c\phi_c}{\epsilon\sigma T_c^4}, \tag{8}$$

where ϵ is the total thermal emissity of the cathode surface.

2.3. *Efficiency characteristics*

To show how the various parameters influence the maximum realizable efficiency η_m as determined from (6) and (8), it is plotted in Figs 3 and 4 for $\epsilon = 0{\cdot}35$ (close to that of tungsten at 3000° K, and a reasonable estimate for composite cathodes at 1200–2000° K) and with j_c determined from

Richardson's equation, with the theoretical value of 120 A cm^{-2} deg^{-2} for A.

Figures 3(*a*) and 4(*a*) are for an anode work function ϕ_a of 1 v, which is about the lowest likely to be achieved in practice. The importance of high current density is at once apparent; this is even more so if ϕ_a cannot be maintained so low, as shown by Figs 3(*b*) and 4(*b*), which are for $\phi_a = 1{\cdot}5$ v. In fact, it is only at current densities of over 1 A/cm^2 that thermionic generators become of much practical interest: this is an important result, since the range of cathode materials which will produce zero-field current densities of this order without deterioration over prolonged periods is quite limited (see, for example, Beck 1959). It is to make this point clear that the presentation of Fig. 3 has been chosen, rather than that of Houston (1959) ((η_m, T_c) for various j_c) or Hernqvist *et al.* (1958) ((η_m, ϕ_c) for various T_c) with which comparison may be made.

The optimum work function at each current density is seen in Fig. 4. In the most useful range for devices likely to have a long life (around 5–10 A/cm^2) the optimum is about 2 v and 2·5 v for $\phi_a = 1$ v and 1·5 v respectively. Thus compound cathodes of the matrix type (see, for example, Beck 1959) hold considerable promise for long-life high-efficiency generators, in spite of the higher Carnot efficiencies of those employing cathodes of higher work function.

Characteristics corresponding to those of Fig. 3(*a*) are shown for several known types of cathode in Fig. 5. In deriving these the 'true work functions' (derived from Richardson's law with $A = 120$ A cm^{-2} deg^{-2}) have been used, rather than work functions derived from 'Richardson plots', since the latter are not strictly relevant, involving as they do the temperature coefficient (Nottingham 1956, Wright 1953). These curves do not correspond exactly with those in Fig. 3(*a*) since account is taken of changes of properties with temperature. The continuous portions represent cathode conditions which might permit reasonably long lives

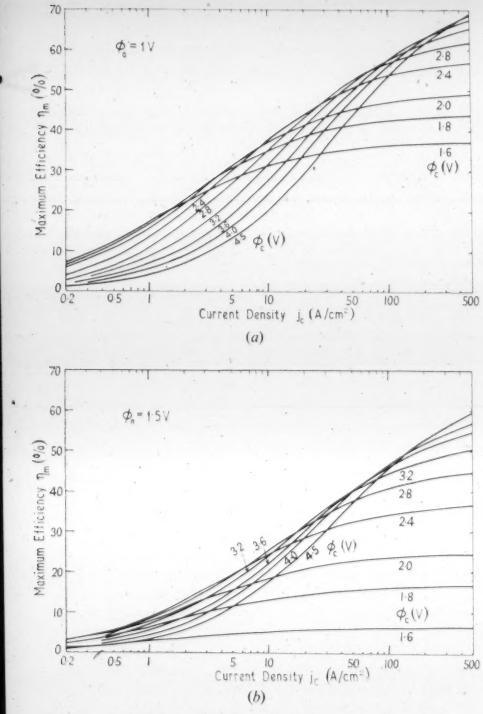


Fig. 3. Calculated maximum efficiencies from ideal diodes, with complete space charge neutralization, cathode radiating as a grey body of emissivity $\epsilon = 0{\cdot}35$.
(*a*) $\phi_a = 1$ v; (*b*) $\phi_a = 1{\cdot}5$ v.

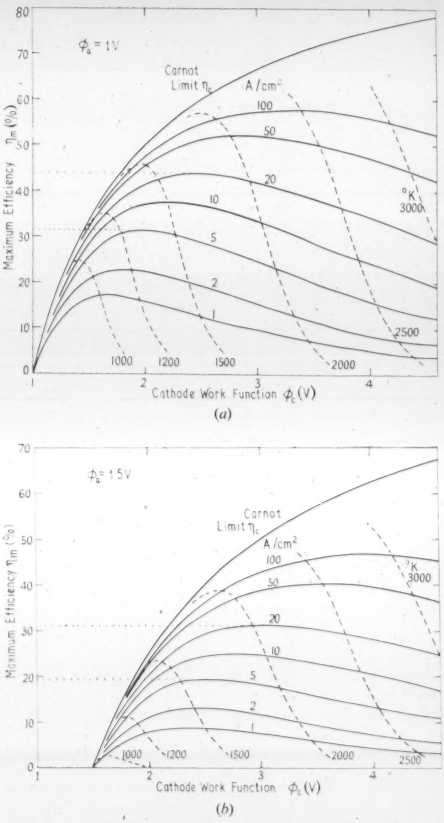


Fig. 4. Maximum efficiencies corresponding to those of Fig. 3 as functions of cathode work function ϕ_c.
(*a*) $\phi_a = 1$ v; (*b*) $\phi_a = 1{\cdot}5$ v.
Broken lines: temperatures in °K.
Dotted lines: maximum efficiencies attainable in unmatched operation (for $j_c = 5$ and 20 A/cm^2)—see section 2.3.

of some thousands of hours or more. Evidently lives in the 10–100 A/cm² region are likely to be very short, although little is known about the lives of the UC–ZrC type (Pidd *et al.* 1959 a,b, Kuczinski 1960, Goldwater and Haddad 1951), the cut-off at 10 A/cm² indicated in Fig. 5 being arbitrary.

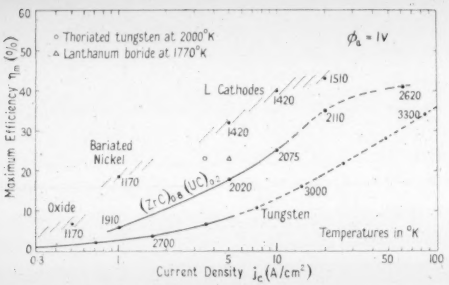


Fig. 5. Estimated maximum efficiencies of generators with existing cathodes, corresponding to those of Fig. 3(*a*), ϕ_a = 1 v. Broken curves indicate regions of heavy evaporation. Single points calculated at maximum usable currents for reasonable life.

Figure 4 also indicates efficiencies attainable at low cathode work functions by operating in unmatched conditions; similar additions can be made to the characteristics derived by Houston and Hernqvist *et al.* These are of little practical interest for the reason given in section 2.1, but they are pointed out since several experiments have been carried out under such conditions, for example those of Hatsopoulos and Kaye (1958), and failure to appreciate this has caused some misunderstandings. (This was drawn to the author's attention by R. V. Harrowell.)

2.4. *Additional factors*

Analyses by Houston (1959) and Rasor (1960) include terms covering the kinetic energy of the electrons and the electrical and thermal conductivity of the lead wire. The first of these is small, but the others lead to contradictory lead wire geometry requirements. Use of the Wiedemann–Franz relation gives an optimum length to diameter ratio. Rasor introduces a figure of merit which is a generalization of (8), and compares it with that commonly used for thermoelectric materials. He shows that these factors lead to a generalization of (6) which is approximated to by

$$\eta_m \simeq 0{\cdot}8\,\eta_c \Big/ \left(1 + \frac{0{\cdot}9}{M}\right) \qquad (9)$$

in typical cases, the numerical factors being insensitive to the actual parameters involved. Hatsopoulos (1960) also compares thermoelectric and thermionic 'figures of merit', but only for the rather artificial 'near equilibrium' case with $T_c \simeq T_a$.

The radiation term P_r depends on the emissivities and reflectivities of the opposing electrode surfaces, and is difficult to estimate with precision. Houston (1959) and Hernqvist *et al.* (1958) restrict their analyses to the case of tunsgten radiating to a black body, but Nottingham (1959b), Hatsopoulos and Kaye (1958), Hatsopoulos *et al.* (1960), and Rittner (1960) use formulae which include reflections between the surfaces.

With actual generators, the effects of residual space charge and the effects of charge interactions and (when a gas is present) collisions must also be taken into account, and some consideration is given to these factors in the following sections. The formulae of this section are useful, however, for assessing and comparing the possible limiting performances of different types.

3. Space charge reduction

Three general methods have been proposed for reducing the space charge of the electrons between the electrodes:

(1) the use of a very close electrode spacing;
(2) the use of combined electric and magnetic fields;
(3) neutralizing the space charge by positive ions.

3.1. *Close-spaced generators*

These consist simply of an anode and cathode arranged so close to one another that no appreciable space charge can build up in the evacuated space between them. A fairly complete theoretical analysis is possible, based on the Langmuir space charge theory (Lindsay and Parker (1959, 1960) give a comprehensive discussion of this and its relevance to thermionic generation). Moss (1957) carried out a fairly detailed analysis, and concluded that a spacing of 0·001 cm or less, with a corresponding degree of flatness and smoothness, is essential for an efficient generator. Hatsopoulos and Kaye (1958) confirmed this experimentally, and obtained an estimated 12–13% efficiency at this spacing (an unfortunate confusion between centimetres and inches occurs in this paper: centimetres are intended, see Nottingham *et al.* 1959).

Successive contributions have been made to the theory, using various approximations, by Webster (1959), Nottingham (1959a), Hatsopoulos *et al.* (1960), Rittner (1960) and Dugan (1960), and it is now in a fairly satisfactory state. The author sees no reason why the exact theory could not now be evaluated, in a form suitable for a digital computer, to give the characteristics of any generator of this type.

Small working generators of this type have been made, using techniques similar to those developed for microwave valve development—see Webster and Beggs (1958) and section 5.

3.2. *Magnetic triodes*

This is the name given to generators making use of combined electric and magnetic fields to overcome the space charge. They have been discussed by Hatsopoulos, Welsh and Langberg (1959) and Welsh *et al.* (1960), and the principle is shown in Fig. 6. The electrons are accelerated by a relatively high potential (~100 v) applied to a third

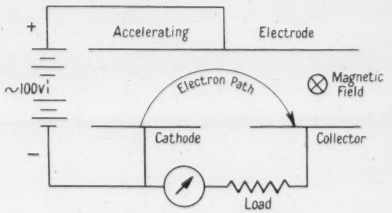


Fig. 6. Principle of the magnetic triode. The magnetic field is normal to the paper.

electrode, corresponding to the anode in an ordinary valve: this overcomes the space charge. A transverse magnetic field diverts them to the collector, as shown.

Additional advantages of this device are the increased possibilities of heat shielding and the possibility of generating a.c. by modulating the magnetic field. Against these, however, are the problems of supplying the latter in the high temperature environment, and the high losses which would be caused by any small electron current reaching the accelerating electrode, because of the high potential.

So far, such devices do not seem to have passed the conjectural stage.

3.3. *Ion-neutralized generators*

In these a gas or vapour is introduced which is ionized to neutralize the electron space charge. Several methods of producing the ions have been considered:

(*a*) by a discharge between the main electrodes;
(*b*) by an auxiliary discharge, involving additional electrodes;
(*c*) by the radiation and particle flux in a nuclear reactor;
(*d*) by 'resonance ionization' of an alkali metal vapour (usually caesium) at a hot surface.

Methods (*a*), (*b*) and (*d*) were listed by Hernqvist *et al.* (1958) who considered the proportion of the generated energy used in producing the necessary ionization, calculating it as an 'equivalent generated voltage loss' (the units are confused in their table I: a loss of 1 v is equivalent to 1 watt used in producing the ions per ampere of generated current). Obviously this loss must be small compared with the generated voltage for an efficient generator.

Method (*a*) is that used by Medicus and Wehner (1951) in their early experiment, using the 'ball of fire' mode of discharge. Hernqvist *et al.* show this to be inherently inefficient, resulting in a loss of about 2 v.

Method (*b*) was found to be nearly as bad in one version, the 'plasmatron' method, in which an auxiliary discharge passes between the main anode and on auxiliary cathode, a loss of about 1 v being estimated. However, work has been carried out subsequently in several laboratories to see whether other forms of auxiliary discharge are more promising.

An ingenious version has been described by Gabor (1961), the essential features of which are illustrated in Fig. 7. This

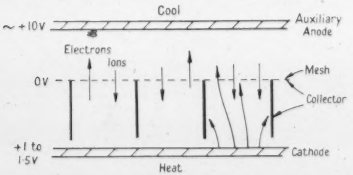


Fig. 7. Principle of Gabor's generator. The region above the mesh contains a luminous ionizing discharge; that in the cells beneath the mesh contains a dark conducting plasma. The electrons flow upwards and the ions downwards, as indicated schematically.

contains argon, or an argon–mercury mixture, at about 1 mm pressure, and the collector consists of metal gauze, of carefully chosen mesh size, through which a proportion of the electrons pass to be accelerated by an auxiliary electrode, causing ionization. The ions are accelerated back through the gauze to neutralize the space charge between the main electrodes. In experiments at low currents the loss has been as low as 0·05–0·1 v. If this is confirmed at higher currents, and if ohmic resistance losses are not too serious, this device would have certain definite advantages over the caesium types: there would be no difficulty in using low temperature cathodes (1300–1700° K) and caesium corrosion problems would be absent. The device described by Gabor incorporates a number of constructional features for increasing the efficiency by increasing the collector area, and for preventing the deleterious effects of the magnetic field of the high currents.

Method (*c*) has been considered by Jablonski *et al.* (1959), who carried out calculations showing that ionization by fission fragments in argon at about 100 mm pressure might be adequate. Jamerson (1960) has followed this up and calculated characteristics which look attractive, but these were not confirmed in a reactor because of design faults. It is still not clear therefore whether it is possible for generators to utilize this principle: at the high gas pressures considered it looks doubtful. A curious feature is that caesium produced as a fission product could lead to a generator of this type changing into one of the caesium type in an hour or two.

Method (*d*) has so far been considered the most promising. The basis is the well-known 'resonance ionization' of alkali metal vapour atoms striking a hot metal surface of work function higher than their ionization potential (see Taylor and Langmuir 1933). The most effective vapour is caesium, and generators in which caesium is ionized in this way have been so extensively studied that they require separate discussion (section 4). Hernqvist *et al.* estimated the loss in generating the ions in this way to be only about 0·01 v. However, several factors are not included in this estimate: all the ions are assumed to arrive at the emitting surface, and any additional radiation losses are neglected. In experiments the losses turn out to be much larger—the results of Wilson (1959) for example correspond to a loss of 6·7 v, which is clearly unsatisfactory. Much of the work at present in progress is aimed at rectifying this.

4. Caesium diodes

4.1. *Experiment*

Caesium vapour introduced into the interelectrode space can perform several functions:

(*a*) provide a source of ions to neutralize the space charge;
(*b*) lower the cathode work function, providing increased electron emission;
(*c*) lower the anode work function, producing higher efficiency.

ϕ_c must, of course, not be lowered unduly or the efficiency will suffer (section 2), and the resonance ionization process responsible for (*a*) will be impaired. However, since $T_a < T_c$, ϕ_a tends to be lowered more than ϕ_c, and by carefully optimizing the temperatures and the caesium pressure, all three effects can be used to advantage. Wilson (1959), using a molybdenum wire cathode at 1900° K and an oxidized silver anode in a caesium pressure of 1·5 mm achieved an estimated efficiency of 9·2% in this way. However, it was evident that the space charge was by no means fully neutralized since, when additional ions were introduced from a second source, the generated current was more than doubled.

Since the resonance ionization process is most efficient with

high cathode work functions and temperatures, experiments have been conducted mainly with refractory metal cathodes operating at 2000–3000° K. Hernqvist, Kanefsky and Norman (1958), Steele (1960a), Block *et al.* (1960) have, among others, all reported the results of experiments using caesium vapour in this range. A very comprehensive investigation is that described by Ranken, Grover and Salmi (1960) (the same team have published a number of previous accounts of this and related work—see Grover *et al.* 1958, Pidd *et al.* 1959b). Fig. 8 shows the characteristics obtained by them using a tantalum emitter and a serrated copper collector for

The possibility of using matrix type cathodes operating at lower temperatures (1300–1500° K), and thus obtaining longer lives, is not entirely precluded by the fact that their average work functions (~1·7 to 2·7 v) are below that (3·88 v) required for 100% resonance ionization of caesium. A certain proportion of the atoms will still be ionized, in accordance with the Langmuir–Saha equation (Langmuir 1925). However, Dobretsov (1960) shows theoretically that sufficient caesium ions are only likely to be produced at high caesium pressures (~1 mm or more). Experimental data so far obtained show that the efficiency of generators with this

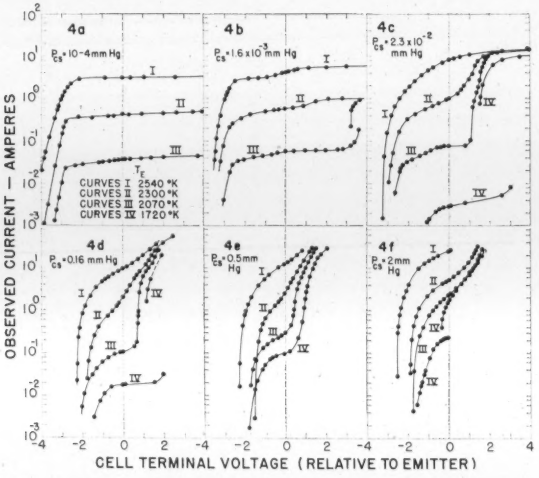


Fig. 8. Characteristics of caesium diode with tantalum emitter and serrated copper collector, observed by Ranken *et al.* (1960). (With acknowledgments.)

various caesium pressures. Comparison with Fig. 1 shows that essentially complete space charge neutralization is achieved at pressures as low as 10^{-4} mm. The current increases as the pressure increases and the cathode work function decreased, particularly at the lower temperatures: this advantage is partly offset by the reduction in the output voltage as a result of the decrease in $\phi_c - \phi_a$, and by collisions in the interelectrode space which introduce an effective ohmic resistance, causing the rounding of the characteristics. The right-hand parts of the characteristics are not relevant: they show the onset of discharge phenomena, which occasionally persist into the retarding region when conditions favour very low voltage discharges. (This account is over-simplified: the original paper gives a much more detailed discussion.)

Clearly there are two principal modes of operation: (*a*) at low caesium pressures (~10^{-4} mm), when there are no collisions in the interelectrode space, and (*b*) at high caesium pressures (~1 mm), when collisions predominate. At high temperatures (over ~1700° K) the latter permits higher current densities, but introduces an effective ohmic resistance. It is, at present, not clear which is likely to prove the most satisfactory (the role of ohmic resistance in limiting the performance of the high pressure type is particularly uncertain) and both are being intensively investigated.

type of cathode is very low at low caesium pressures, but increases rapidly as the pressure is raised: the limit is not yet known.

The possibility of making use of the patchy nature of the surface of this type of cathode has been considered, for example by Nottingham (1959b), and is being investigated experimentally. There is a danger that ions produced at regions of high work function may be trapped in space charge 'pockets', and it may be necessary to bias such areas a volt or two positive to the low work function areas to overcome this: this would add complication, but does not seem impossible. Dobretsov (1960) calculates that a patchy cathode is no better than a uniform one with a work function equal to its average, but this part of his analysis is not very convincing. Morgulis and Naumovets (1960) obtained about 0·5 w/cm² at an estimated 4% efficiency using a matrix cathode at 1300° C, with an estimated 200 hr life, which they interpret in terms of ionization at uncoated areas of tungsten on the cathode.

An arrangement akin to this in which a third electrode is added, operating at the cathode temperature, but with a higher work function, to ionize the caesium has been announced by Hernqvist (1960). An estimated efficiency of 14% is claimed with the cathode at 1100° C.

4.2. *Theory*

Theoretically, knowledge of conditions in the interelectrode space is far from complete at present. As soon as there is appreciable ionization it is evident that a relatively neutral plasma will form, with space charge sheaths at the electrodes, the potential distribution changing in some such manner as shown in Fig. 9 as the ionization is increased. Evidence for

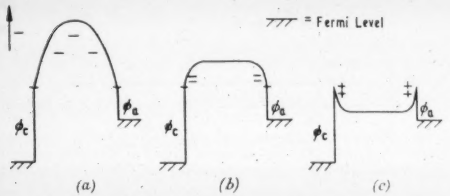


Fig. 9. Conjectural potential diagrams showing electron space charge (*a*) in vacuum, (*b*) partially neutralized, (*c*) completely neutralized, by positive ions. Signs of residual changes as indicated.

the formation of such sheaths has been given by Steele (1960b), for example, and they have been discussed by Nottingham (1960). With a positive ion sheath at the cathode, as in Fig. 9(*c*), a positive ion current density $j_i \geqslant j_e(m_p/m_e)^{1/2}$ is required (Langmuir 1929) to permit an electron current j_e to flow. If the available ion current is smaller than this the distribution will revert to the form of Fig. 9(*b*), and the electron current will be reduced accordingly.

At low pressures collisions can be ignored, and some progress has been made in modifying the vacuum diode space-charge theory to include the effect of ions, notably by Auer and Hurwitz (1959), Auer (1960) and Franklin (1960). The first two find instabilities in their solutions, which may give rise to oscillations, and these are also shown to arise in certain conditions by Eichenbaum and Hernqvist (1961), but Franklin doubts whether this type of instability occurs in reality. (Oscillations have been frequently observed, particularly at near zero anode potentials, for example by Johnson (1960) and Garvin *et al.* (1960), and their elucidation, and possible control to produce a.c. (see *Electronics* (1960a)) is currently receiving much attention).

Nottingham (1959b) gives a detailed treatment of the ion production process, including patchy cathodes, and shows that a low pressure diode may be regarded as equivalent to a vacuum diode with a much closer spacing. The potential in the cathode sheath is evaluated by Enoch and Rankin (1959) and Nottingham (1960): for a typical case the former obtains a sheath thickness of the order of 10^{-3} cm.

At high pressures conditions are much more like those in ordinary discharge plasmas, the motion of the ions and electrons more or less randomized by collisions, and the currents to and from the electrodes largely determined by thin space-charge sheaths.

The most detailed treatment is that of Lewis and Reitz (1959a, b, 1960), who develop an analogy with the thermoelectric effect, calling a generator of this type a 'plasma thermocouple'. In the present author's opinion this is unfortunate, and only serves to obscure the situation: conditions in the thermionic case are so far from thermodynamic equilibrium that the analogies with the Seebeck coefficient, thermoelectric power, etc., are rather forced, and of dubious utility. The later contributions by these authors draw progressively less on these concepts, and form a considerable contribution to the understanding of the subject, taking into account the thermal and electrical conductivity of the plasma, and the electron and ion temperature distribution as a result of ohmic heating. They assume, however, that the ionization occurs in the interelectrode space (as a result of electron collisions, thermal ionization and photoionization) and ignore ion production at the cathode surface. This is a serious shortcoming in view of the experimental evidence showing the importance of the cathode surface. Nottingham (1960) takes into account both surface ionization and electron collision ionization, but his theory has not yet been evaluated sufficiently to show what the relative importance of the latter is likely to be.

The present author feels that considerable insight will be obtained by developing further the treatment of Lewis and Reitz, dropping the thermocouple analogy entirely as a logical sequel to the trend already shown, and systematically applying the methods of analysis familiar in discharge theory (for example, Malter *et al.* 1951, 1952, 1955, Cayless 1958).

5. Working generators

Most of the experimental results quoted in earlier sections have been obtained from laboratory devices, usually in glass bulbs, and with electrically heated cathodes. It is important to note that the efficiencies of these have usually been estimated from the *measured* electrical characteristics, and some form of *estimated* heat losses from the cathode, using various assumptions, sometimes optimistic. Those quoted in this article are thought by the author to be reasonable estimates.

Some generators which have been constructed and which actually *generate* electricity from non-electric heat sources will now be described.

A caesium generator of tubular form suitable for operating from rocket exhaust gases has been described by Block *et al.* (1960). This consists of a molybdenum or tantalum tube which forms a cathode of 150 cm^2 area, surrounded by a copper tube which forms the anode. With the cathode heated to 2250° C, the anode at 150–200° C, and a caesium pressure of 0·01 to 0·1 mm, this generated 269 W (164 A at 1·64 V across 0·01 Ω) with a thermal input of 10·1 kW—an efficiency of 2·7%. With a weight of 3½ lb this represents 80 W/lb, a useful amount of power in a rocket with a virtually unlimited heat supply. The tests were carried out in vacuum or an inert atmosphere: life in use in a rocket would be short.

Grover (1959) describes a generator being developed for use in a nuclear rocket reactor. In this a cylinder of UC : ZrC (the uranium enriched to 94%) is used as both the cathode and the heat source. In an experimental version this was about 1 in. long × ¼ in. in diameter, positioned inside a stainless steel cylinder which acted as anode; an oil coolant system controlled the anode temperature and caesium pressure. In the reactor a neutron flux of about 10^{13} raised the cathode to about 2000° C, producing 3·8 V (open circuit) and 38 A (short circuit). The power could not be measured under load, but was estimated to be about 30 W. Radiochemical analysis of the cathode after removal enabled a rough estimate to be made of the efficiency of about 5%. A rather similar generator described in *Electronics* (1960c) produced 90 W (21 W/cm^2) at an estimated 10%, at 1930° C.

Small disk type vacuum and caesium generators are described in *Electronic Industries* (1960). Both consist essentially of cathode and anode disks sealed together by a ceramic ring, much as in Fig. 2. The vacuum type is stated

to be 'in production': it is about 1 in. in diameter and produces about 1 w at over 2·5% efficiency at about 1000° c cathode temperature. This is achieved by constructional techniques which enable the two optically flat electrode surfaces to be kept precisely 0·0005 in. apart over prolonged periods at the operating temperatures. The caesium generator is stated to be 'production type'. Similar to the other, but about 1½ in. in diameter, and with appendages containing the caesium, this has produced up to 2 w/cm² at about 1300° c cathode temperature, with an efficiency estimated to be between 8 and 12%.

A prototype of a solar-powered generator described by Leovic and Mueller (1960) had a tantalum disk cathode on which the radiation was focused, and a silver coated anode with a work function reduced to 1·8 v by the caesium. In tests using an arc image furnace nearly 2 w/cm² were generated with cathode and anode at 2480 and 700° k respectively, and caesium at about 2 mm. A similar, but more completely engineered, device described by Oman and Street (1960) had not been fully tested, but 24 w (8 w/cm²) at a cathode temperature of 2000° k were measured using ion bombardment heating, giving an estimated efficiency of about 10%.

6. Applications

The greatest stimulus for the development of thermionic generators in the United States has been the potential space applications. In rockets, where power is needed to operate auxiliary equipment, the high power densities and power/weight ratios (several w/cm² at only a few g/w) already attained by the high temperature (>2000° k) devices, together with structural rigidity and absence of moving parts, provide clear advantages over other types of generator. Long life, and even efficiency, are relatively unimportant.

For satellites, the high power density (compared with the 10 mw/cm² or so attainable from photovoltaic solar cells), low weight and high anode temperature (which assists in the disposal of waste heat by radiation) are advantageous, together with the low cost compared with that of solar cells. The necessity to concentrate sunlight by accurately oriented mirrors is a disadvantage, and for this reason, together with the long life required, generators operating at low cathode temperatures (below 1700° k) are preferable.

The efficiency attainable, 2–10% in currently working devices, 10–20% in laboratory devices and 20–30% forseeable as feasible, is too small for thermionic generators to be attractive as direct alternatives to conventional steam turbine–dynamo plant for large scale terrestrial power generation. The overall efficiency of the conventional process is approaching 40–45% for combustion heating and 20–35% for nuclear heating in large stations. However, these efficiencies are likely to improve further only very slowly because of the Carnot limitation and the difficulty of devising materials which would permit the operation of rotating machinery at higher temperatures.

By combining conventional and thermionic generators, however, this barrier could be overcome and useful increases in overall efficiency obtained. This is the process known as thermodynamic 'topping': the thermionic generator takes in heat at, say, 1100–1300° c, rejecting it at 400–600° c for use by turbines. Fig. 10 illustrates the theoretical improvement in overall efficiency obtainable in this way: for example, a thermionic generator of 15% efficiency could increase the overall efficiency of a steam plant from 30% to over 40%, or from 40% to nearly 50%. Economically this appears to be a sound proposition, which would justify a fairly high capital cost for the additional equipment. It would also seem feasible to use thermionic generators for providing supplementary power at peak periods. The development of generators operating at cathode temperatures in the range 1100 to 1300° c is desirable for these applications because of the long life required, and also because higher temperatures are not readily available, at least from non-nuclear fuels.

For nuclear power sources, the generator could form part of the reactor core, as in the example described in section 5, and with reactor development higher cathode temperatures may become feasible. In this case, it is interesting to note that the electron cooling provides a highly efficient heat-transfer mechanism from the fuel-element cathode to the surrounding anode can—the generation is almost a bonus! (Wigner 1959).

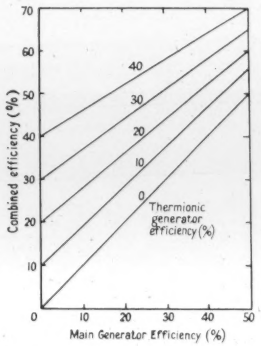


Fig. 10. Possible improvement in efficiency obtainable by combining thermionic generators with conventional generating plant ('topping').

An attractive alternative is an entirely direct-conversion station, using combined thermionic and thermoelectric generators, the anode heat from the former being used by the latter. Several such devices have been proposed, and at least two actually constructed (Celent 1960, *Electronics* 1960d).

Other possible applications are for small, quiet, remote or unattended generators, or where the direct generation of low voltage d.c. would be advantageous, as in electrolytic plant or certain traction applications. In many of these cases the choice might lie between thermionic generators and other direct conversion devices, such as thermoelectric or photovoltaic generators, or fuel cells. For further discussion of applications the references given at the end of section 1 may be consulted.

7. Further developments and problems

Although considerable success has been achieved in the construction of generators which produce useful amounts of power, there remain many problems, some of a formidable nature, to be solved before they are likely to find widespread utility. Most of those constructed so far operate satisfactorily only at temperatures over 2000° c, and to obtain long lives very substantial developments in materials are required. Even in the 1100–1300° c range it is by no means easy to devise constructions which will withstand prolonged operation in the oxidizing or corrosive atmospheres produced

in ordinary combustion processes, and the attack of materials by caesium vapour is a serious problem. In nuclear reactors the situation is, in one respect, easier since only the cathode itself need be heated to the high temperature, as described in section 5, but problems of compatibility, absorption cross sections, lead design, radiation damage and fission product disposal arise which are even more formidable.

At low cathode temperatures, although the constructional and materials problems are easier, provision of adequate space-charge neutralization is more difficult, and generators using third electrodes or auxiliary discharges may be more suitable than plain diodes. It is difficult to see how the vacuum close-spaced type can be scaled up very much from the very small models made so far. The development of cathodes with high zero-field emissions is a problem, particularly when the work function is required to be as high as possible at the same time, thus requiring 'anomalous' emission properties. Most available high current density cathodes only produce their high currents under high fields, often under pulsed conditions.

The development of anodes with the lowest possible work functions present a challenge to ingenuity. For the high temperature generators, using metal cathodes, caesium films with $\phi \simeq 1{\cdot}8$ v are adequate but, as section 2 shows, work functions of 1–1·5 v are highly desirable, especially when lower temperature cathodes are used. Photo-cathodes of the Cs–O–Ag type come to mind, but these are normally unstable at high temperatures. At high gas pressures the resistance of the plasma is probably a limiting feature. Even at low pressures, high electron densities may lead to the electron–electron interaction becoming important. Lewis and Reitz (1960) and Gabor (1961) have considered this and shown that it is appreciable, but little is known about the limitations which it may impose on the maximum power attainable.

The self magnetic field of the electron current provides an important limitation, and provision must be made for countering its effects in any generator with a high current flow. Schock (1960) considers this in detail and shows that its effects may be ameliorated by superimposing an additional magnetic field parallel to the direction of current flow. Gabor (1961) describes a design aimed at minimizing the effect.

There is considerable scope for ingenuity in designing simple, robust constructions which optimize the flow of the high thermal and electric currents involved, particularly when arranging for units to be cascaded electrically to produce higher output potentials (see, for example, the method for cascading in a reactor proposed by Salmi (1960)). Thermal cascading is also possible, the output heat from a high temperature generator being used to operate a low temperature generator. Mention has been made in section 5 of combining thermionic and thermoelectric generators in this way. Two thermionic generators could be arranged similarly, and this has been tried by Saldi (1960), but his results are inconclusive: calculations made in the author's laboratory suggest that there is, in fact, not much to be gained in this case.

Heat shields between the electrodes would increase efficiency, and the experiments of Ranken *et al.* (1960) show that this is feasible, at any rate in generators with high caesium pressures, without much detriment to the electrical performance. Development of electrode surfaces of low thermal emissivity and reflectivity, combined with the other properties required, is a parallel problem.

Although invertor circuits and devices are now fairly common and efficient, a generator producing a.c. directly would have obvious advantages. Those in which ionization is effected by means of a third electrode or auxiliary discharge can probably all be arranged to produce a.c. by using a suitable feedback system to reduce the ionization as the current develops, though with some loss in efficiency, since the radiation loss will continue during the 'off' half cycle: possibly a 'full-wave' generator with two collectors could be devised to overcome this. In the case of diodes, the oscillations, which have received so much attention (see section 4.2), might prove a possible source of a.c. if they can be controlled.

8. Conclusions

Thermionic generators, which first came into prominence only some three years ago, have now firmly established themselves as potentially useful generators of electricity directly from heat. High power densities, high power/weight ratios and high operating temperatures, giving a high limiting Carnot efficiency, are their principal virtues. These give them clear advantages for certain applications over their nearest rivals, thermoelectric generators, in the present state of development, although spectacular improvements in thermoelectric materials could largely nullify some of these if they could be made.

Devices producing several w/cm^2 at 2–10% efficiency have been constructed and demonstrated, and laboratory experiments have yielded 10–20 w/cm^2 at 10–20% efficiency. Theoretical assessments indicate that, even allowing for practical limitations, over 20 w/cm^2 at 20–30% efficiency should be ultimately attainable.

Although several applications, future developments and problems have been described, many more will readily occur to anyone engaged in this subject. The whole field presents a fascinating and fertile ground for research and development, and there is no reason to suppose that the present intense activity will do anything but increase.

Acknowledgments

The author thanks his colleagues at the Research Laboratory of A.E.I. (Rugby) Ltd. and elsewhere with whom he has discussed this subject on many occasions, particularly Mr. J. A. Thorp, who also prepared the data on which Fig. 5 is based.

References

AUER, P. L., 1960, *J. Appl. Phys.*, **31**, 2096.
AUER, P. L., and HURWITZ, H., 1959, *J. Appl. Phys.*, **30**, 161.
BECK, A. H. W., 1959, *Proc. Instn Elect. Engng*, **106**, *Pt B*, 372.
BLOCK, F. G., CORREGAN, F. H., EASTMAN, G. Y., FENDLEY, J. R., HERNQVIST, K. G., and HILLS, E. J., 1960, *Proc. Inst. Radio Engrs, N.Y.*, **48**, 1846.
CAYLESS, M. A., 1958, *J. Electron. and Control*, **4**, 237.
CELENT, C. M., 1960, *Electronic Industries*, **19** (9), 102.
CHAMPEIX, R., 1951, *Le Vide*, No. 31, 936.
DOBRETSOV, L. N., 1960, *Soviet Phys. Tech. Phys.*, **5**, 343.
DUGAN, A. F., 1960, *J. Appl. Phys.*, **31**, 1397.
EGLI, P. H., 1961 (Ed.), *Thermoelectricity* (New York: John Wiley).
EICHENBAUM, A. L., and HERNQVIST, K. G., 1961, *J. Appl. Phys.*, **32**, 16.
Electronic Industries, 1960, **19**, (9), 117.
Electronics, 1960a, **33**, (5), 78 (29th Jan.).
—— 1960b, **33**, (11), 162 (11th March).
—— 1960c, **33**, (49), 78 (2nd Dec.).
—— 1960d, **33**, (51), 112 (16th Dec).

ENOCH, J., and RANKEN, W. A., 1959, *Proc. 4th Int. Conf. on Ionization Phenomena in Gases, Uppsala*, **1,** 482 (Amsterdam: North Holland).
FEASTER, G. R., 1958, *J. Electron. Control*, **5,** 142.
FRANKLIN, R. N., 1960, *J. Electron. Control*, **9,** 385.
GABOR, D., 1961, *Nature, Lond.*, **189,** 868.
GARVIN, H. L., TEUTSCH, W. B., and PIDD, R. W., 1960, *J. Appl. Phys.*, **31,** 1508.
GOLDSMID, H. J., 1960, *Brit. J. Appl. Phys.*, **11,** 209.
GOLDWATER, D. L., and HADDAD, R. E., 1951, *J. Appl. Phys.*, **22,** 70.
GRATTRIDGE, W., 1960, *Inst. Radio Engrs Convention Record, Pt* 9, 54.
GROVER, G. M., 1959, *Nucleonics*, **17,** 54.
GROVER, G. M., ROEHLING, D. J., SALMI, E. W., PIDD, R. W., 1958, *J. Appl. Phys.*, **29,** 1611.
HARVARD, 1959, *Thermoelectricity: A Report for Business* (Boston: Harvard Business School).
HATSOPOULOS, G. N., 1960, in KAYE and WELSH, 1960, Chapter 3.
HATSOPOULOS, G. N., and KAYE, J., 1958, *Proc. Inst. Radio Engrs, N.Y.*, **46,** 1374 (reprinted in KAYE and WELSH, 1960).
HATSOPOULOS, G. N., KAYE, J., and LANGBERG, E., 1960, *Inst. Radio Engrs, Trans Electron Devices*, **7,** 117 (reprinted in KAYE and WELSH, 1960).
HATSOPOULOS, G. N., WELSH, J., and LANGBERG, E., 1959, *Electronics*, **32,** (46), 69 (13th Nov.).
HERNQVIST, K. G., 1960, *Electronic News*, **5** (225), 4 (3rd Oct.).
HERNQVIST, K. G., KANEFSKY, M., and NORMAN, F. H., 1958, *R.C.A. Rev.*, **19,** 244 (reprinted in KAYE and WELSH, 1960).
HOUSTON, J. M., 1959, *J. Appl. Phys.*, **30,** 481.
HURST, H., 1958, *Brit. Patent Spec. No.* 797, 872.
JABLONSKI, F. E., LAFFERT, C. B., SILVER, R., HILL, R. F., and LOUGHRIDGE, D. H., 1959, *J. Appl. Phys.*, **30,** 2017.
JAMERSON, F. E., 1960, *Inst. Radio Engrs Internat. Convention Record, Pt* 9, 66.
JAUMOT, F. E., 1958, *Proc. Inst. Radio Engrs, N.Y.*, **46,** 538.
JOHNSON, F. M., 1960, *Rep. 20th Ann. Conf. Phys. Electronics, M.I.T.*, p. 88 (Cambridge, Mass.: M.I.T.).
KAYE, J., and WELSH, J. A., 1960, (Ed.), *Direct Conversion of Heat to Electricity* (New York: John Wiley).
KUCZINSKI, G. C., 1960, *J. Appl. Phys.*, **31,** 1500.
LANGMUIR, I., 1925, *Proc. Roy. Soc. A*, **107,** 61.
—— 1929, *Phys. Rev.*, **33,** 954.
LEOVIC, W. J., and MUELLER, M. W., 1960, *Elect. Engng*, **79,** 979.
LEWIS, H. W., and REITZ, J. R., 1959a, *J. Appl. Phys.*, **30,** 1439.
—— 1959b, *J. Appl. Phys.*, **30,** 1838.
—— 1960, *J. Appl. Phys.*, **31,** 723.
LINDSAY, P. A., and PARKER, F. W., 1959, *J. Electron. Control*, **7,** 289.
—— 1960, *J. Electron. Control*, **9,** 81.
MALTER, L., JOHNSON, E. O., and WEBSTER, W. M., 1951, *R.C.A. Rev.*, **12,** 415.
—— 1952, *R.C.A. Rev.*, **13,** 163.
—— 1955, *R.C.A. Rev.*, **16,** 82, 498.
MEDICUS, G., and WEHNER, G., 1951, *J. Appl. Phys.*, **22,** 1389.
MORGULIS, N. D., and NAUMOVETS, A. G., 1960, *Soviet Phys. Solid State*, **2,** 501.
MOSS, H., 1957, *J. Electronics*, **2,** 305.
NOTTINGHAM, W. B., 1956, *Handbuch der Physik*, **22,** 1 (Berlin: Springer-Verlag).
—— 1959a, *J. Appl. Phys.*, **30,** 413.
—— 1959b, *Proc. 4th Int. Conf. on Ionization Phenomena in Gases, Uppsala*, **1,** 486 (Amsterdam: North Holland) (reprinted in KAYE and WELSH, 1960).
—— 1960, *Rep. 20th Ann. Conf. Phys. Electronics, M.I.T.*, p. 95 (Cambridge, Mass.: M.I.T.).
NOTTINGHAM, W. B., HATSOPOULOS, G. N., and KAYE, J., 1959, *J. Appl. Phys.*, **30,** 440.
OMAN, H., and STREET, G., 1960, *Elect. Engng*, **79,** 967.
PIDD, R. W., GROVER, G. M., ROEHLING, D. J., SALMI, E. W., FARR, J. D., KRIKORIAN, N. H., and WITTEMAN, W. G., 1959a, *J. Appl. Phys.*, **30,** 1575.
PIDD, R. W., GROVER, G. M., SALMI, E. W., ROEHLING, D. J., and ERICKSON, G. F., 1959b, *J. Appl. Phys.*, **30,** 1861 (reprinted in KAYE and WELSH, 1960).
RANKEN, W. A., GROVER, G. M., and SALMI, E. W., 1960, *J. Appl. Phys.*, **31,** 2140.
RASOR, N. S., 1960, *J. Appl. Phys.*, **31,** 163.
RITTNER, E. S., 1960, *J. Appl. Phys.*, **31,** 1065.
SALDI, I. T., 1960, *Inst. Radio Engrs Convention Record, Pt* 3, 3.
SALMI, E. W., 1960, *J. Electrochem. Soc.*, **107,** 1013.
SCHLICHTER, W., 1915, *Dissertation*, Göttingen.
SCHOCK, A., 1960, *J. Appl. Phys.*, **31,** 1978.
SPRING, K. H., 1961, *Nature, Lond.*, **190,** 297.
STEELE, H., 1960a, in KAYE and WELSH, 1960, Chap. II.
—— 1960b, *Rep. 20th Ann. Conf. Phys. Electronics, M.I.T.* (Cambridge, Mass.: M.I.T.), p. 76.
TAYLOR, J. B., and LANGMUIR, I., 1933, *Phys. Rev.*, **44,** 423.
WEBSTER, H. F., 1959, *J. Appl. Phys.*, **30,** 488.
WEBSTER, H. F., and BEGGS, J. E., 1958, *Bull. Amer. Phys. Soc.*, Ser. II, **3,** 266.
WELSH, J. A., HATSOPOULOS, G. N., and KAYE, J., 1960, in KAYE and WELSH, 1960, Chap. 5.
WIGNER, E. P., 1959, *Nuclear Sci. and Eng.*, **6,** 420.
WILSON, V. C., 1959, *J. Appl. Phys.*, **30,** 475.
—— 1960, in KAYE and WELSH, 1960, Chap. 7.
WRIGHT, D. A., 1953, *Proc. Instn Elect. Engrs*, **100,** Pt III, 125.
YAFFEE, M., 1959, *Aviation Week*, 23*rd November*, p. 92.

Note added in proof.—The following relevant references have appeared since this article was prepared:—

CARABATEAS, E. N., PEZARIS, S. D., and HATSOPOULOS, G. N., 1961, *J. Appl. Phys.*, **32,** 352.
Electronics, 1961, **34,** 94 (14th April).
Electronics Weekly, 1961, No. 32, 19 (12th April).
HERNQVIST, K. G., 1961, *R.C.A. Rev.*, **22,** 7.
HOWARD, R. C., 1961, *J. Instn. Elec. Engrs.*, **7,** 284.
INGOLD, J. H., 1961, *J. Appl. Phys.*, **32,** 769.
JOHNSON, F. M., 1961, *R.C.A. Rev.*, **22,** 21.
ZOLLWEG, R. J., and GOTTLIEB, M., 1961, *J. Appl. Phys.*, **32,** 890.

Investigation of the electrical contact properties of granular carbon aggregates

by E. D. MACKLEN, M.Sc., Ph.D., Standard Telecommunication Laboratories Ltd., Harlow, Essex

Paper presented on 7th April 1961, at The Institute of Physics and The Physical Society Symposium on Electrical Contacts

MS. received 14th April 1961

Abstract

The heat treatment of granular carbon aggregates in various ambient gases has been studied. In inert gases (hydrogen, nitrogen, argon, etc.) and in vacuo, *either a constant or a decreasing electrical resistance was observed for increasing heat treatment temperature. Ambient gases containing either free or combined oxygen produced a sharp resistance maximum at about 530° C. The increase in resistance is attributed to the formation of a non-conducting surface oxide layer which decomposes on increasing the heat treatment temperature.*

Introduction

GRANULAR carbon has an important use in the telecommunication industry in the carbon microphone. A combination of properties, so far found only in partially carbonized anthracite, has made carbon eminently suitable for this application. With time and use, the electrical properties of the carbon aggregate alter, causing deterioration in the microphone performance. This deterioration usually shows itself by an increase in resistance and noise level, together with a decrease in the modulation efficiency of the carbon aggregate. Several causes of deterioration have been suggested, but few have been thoroughly investigated; the majority of the past work has been concerned with the production of a stabilized carbon by mechanical treatment.

During normal use of the telephone, the feed current of the transmitter capsule causes localized heating of the carbon–carbon contacts to temperatures possibly in excess of 1000° C. These conditions are therefore suitable for rapid desorption and adsorption of ambient gases to occur together with localized surface oxidation within the contact area. In the present work an investigation has been made of the effects on the permanent electrical properties of heating the carbon aggregate in various gases, in particular those gases normally encountered by a transmitter capsule.

Little previous work on this subject has been found in the literature. Siebel (1921) showed that carbon filaments had a 4–6% higher resistance in a gas atmosphere than *in vacuo*, the order of resistances being: ammonia > sulphur dioxide > carbon dioxide > air > vacuum. A few years later, Wright and Marshall (1928) measured the resistance of a single contact *in vacuo* before and after heat treating at 1700° C. The heat treatment was found to lower the contact resistance, which could be restored by re-exposing to the air.

Waetzmann and Gigling (1928) made the first study of aggregates in various atmospheres, and showed that a microphone had a resistance of 86 ohms in damp air, 120 ohms in dry air, and 136 ohms in a vacuum of 10^{-1} mm Hg. Powdered coal, pressed into sticks and then carbonized in nitrogen, was found by Sandor (1944) to give a 60% increase in resistance after 23 days exposure to air. Alcohol and water vapour were also found to affect the electrical resistance. The contact resistance of a synthetic carbon powder, heat treated in air between 200° and 900° C, was found by Hirabayashi and Toyoda (1954) to pass through a maximum at 500–600° C. This maximum was attributed to surface oxidation.

Experimental

Materials

Since the primary object of this study was to investigate the causes of deterioration in carbon granules of the type used in microphones, commercially available microphone material obtained by carbonizing anthracite and sieving to BSS 60/80 mesh sizes was used throughout. The commercial gases supplied by The British Oxygen Company were considered sufficiently pure for these experiments, and they were dried by passage through anhydrone and stored in glass reservoirs.

Apparatus and procedure

The carbon was placed in a Pyrex–silica tube fitted with a conductivity cell for resistance measurements at the Pyrex end and sealed off at the silica end. A side arm in the Pyrex section connected the tube, via a ball joint, to a vacuum apparatus with manometers, gas reservoirs, and vacuum pump. A diagram of this tube is shown in Fig. 1.

The resistance of 3 g of granular carbon was measured in the conductivity cell by passing 10 mA from a constant current source through the carbon, and tapping the cell gently until a minimum value of the voltage across the electrodes was obtained. The cell electrodes were rhodium plated to minimize and stabilize carbon–electrode contact resistance. Previous experiments over a range of cell currents had shown that 10 mA gave a reproducible value for the resistance which was representative of the carbon. After measuring the resistance of the carbon in air in the cell, the tube was rotated about the ball joint and the carbon tipped into the silica section of the tube. With the tube horizontal to expose maximum carbon surface, the air was evacuated for several minutes, the carbon returned to the conductivity cell, and the minimum resistance under a rough vacuum (*c.* 10^{-2} mm Hg) measured as before. With the carbon again spread out, the ambient gas was admitted at approximately atmospheric pressure, and after a few minutes the resistance was remeasured. The heat treatment of the carbon was carried out by inserting the silica end of the tube containing the carbon into a horizontal tube furnace, and measuring the temperature by a thermocouple placed in contact with the silica tube beneath the carbon. After about

half an hour at the selected temperature the furnace was removed, the temperature allowed to fall to room temperature (20–22° c) and the carbon resistance measured in the 'stale' gas, in vacuum, and then with fresh ambient gas

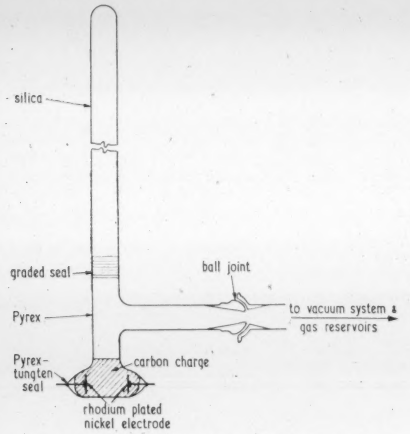


Fig. 1. Pyrex–silica tube with conductivity cell.

admitted. The shapes of the resistance–temperature plots were the same for the three sets of values, although the vacuum resistance was usually slightly higher than the other two. The vacuum resistance measurements have been reported here, for they were felt to be more representative when comparing the effects on the carbon of the heat treatments in various ambients.

A second conductivity cell was constructed for measuring the resistance of carbon granules under normal room temperature conditions. The cell was of a lamellar construction with alternate vertical layers of conductors (rhodium-plated brass plates $\frac{1}{32}$ in. thick) and insulators (Tufnol sheet $\frac{1}{8}$ in. thick). A cavity was cut from this structure to receive the granular carbon, as shown in Fig. 2. In this cell the three inner electrodes stopped short by a $\frac{1}{4}$ in. of the top of the insulators. A standard volume (1·5 cm^3) of carbon was

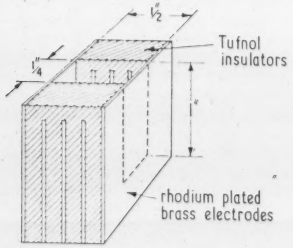


Fig. 2. Lamellar conductivity cell.

dropped into the cell by means of a hopper so that the upper surface of the carbon aggregate was above the inner electrodes. A feed current from a constant d.c. source was passed between the outer two electrodes, and the potential difference measured between any two (usually the outer pair) of the inner electrodes to give the carbon resistance. This arrangement of probe electrodes eliminated the carbon–electrode contact resistance at the current electrodes. Using this technique, rapid resistance values for a particular carbon were obtained reproducible to within about ±3%.

The resistance obtained in this manner was equal in value to that obtained by tapping the cell to a minimum value, but it was obtained much more rapidly.

Results and discussion

The experimental values for argon have been inserted in Fig. 3 to indicate the small amount of scatter experienced in this investigation. The actual values for the other ambient gases have been omitted so that the general trends of the plots may be clearly seen. These values are to be found in Table 1.

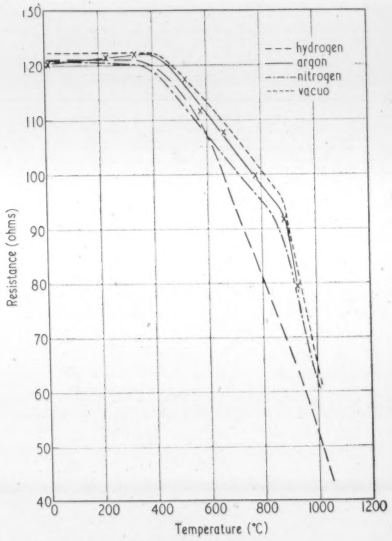


Fig. 3. Heat treatment in H_2, N_2, A and *vacuo*.

From a comparison of the effects on electrical resistance of heating the granular carbon in various ambient gases, shown in Figs 3 and 4, it is immediately apparent that there is a marked difference between the two groups of gases. Heat treatment in oxygen and carbon dioxide leads to a sharp resistance peak at about 530° c, whereas heating in gases not containing oxygen has little effect up to about 400° c. Above this temperature, the resistance decreases steadily until about 900° c. Fig. 3 shows that a more rapid decrease in resistance with temperature occurs above 900° c. Since resistance increases were caused by both oxidizing gases, while with reducing or inert gases and vacuum no increase was observed, an explanation of this phenomenon was sought in the literature of carbon oxidation.

The chemical reaction between carbon and oxidizing gases

Table 1. *Carbon aggregate resistances after heat treatment in various ambient gases*

Vacuum

Temperature (° C)	22	52	135	183	276	333	420	497	635	810	860	988
Resistance (Ω)	122	123	124	122	122	123	121	119	112	100	92	66
Temperature (° C)	22	302	382	468	538	817	860	932	972	1020		
Resistance (Ω)	123	123	122	119	115	99	96	82	72	61		

Nitrogen

Temperature (° C)	23	82	159	217	287	388	495	578	707	923		
Resistance (Ω)	120	121	121	120	119	120	113	109	99	94		
Temperature (° C)	360	459	560	618	777	852	958	1006				
Resistance (Ω)	118	117	109	107	95	90	70	60				

Hydrogen

Temperature (° C)	46	140	192	242	318	400	508	602	692	742	885	958
Resistance (Ω)	122	120	121	120	122	119	117	105	95	89	69	57
Temperature (° C)	267	350	459	512	604	686	777	856	966	1017		
Resistance (Ω)	120	119	118	115	108	94	83	75	58	50		

has been very fully investigated over a wide range of temperatures, pressures, and flow rates. From all of this work one very important fact has emerged, namely that the first stage in the oxidation of carbon involves the formation of an oxygen complex on the surface of the carbon (Smith 1959). Once an equilibrium has been established, the oxygen content of a carbon heated in oxygen or carbon dioxide has been found (Gulbransen and Andrews 1953, Puri *et al.* 1958) to depend primarily on the heat treatment temperature rather than the heating period. For a carbon black oxidized in air (Snow *et al.* 1958) the oxygen content increased rapidly from 3% at an oxidation temperature of 200° C to a maximum value of 11% at 460° C, then fell to 3% at 700° C, continuing to fall to 1% at 1100° C.

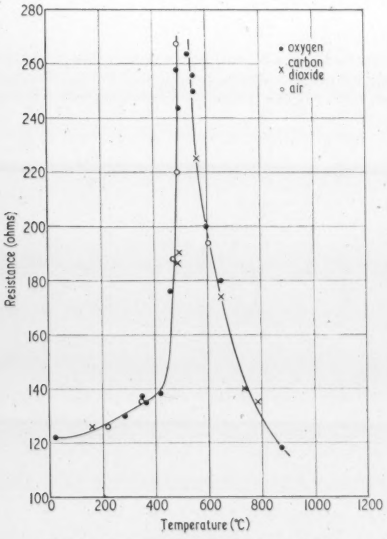


Fig. 4. Heat treatment in air, O_2, and CO_2.

The general shapes of this oxygen-content curve and the contact resistance–temperature curve are the same, suggesting that the increase in contact resistance is due to the formation of a non-conducting oxygen complex on the surface of the granular carbon. The exact position of the maximum in oxygen content depends on the rates of formation and decomposition of the oxygen complex. Since both of these rates will be affected by the nature of the carbon, impurity content, and surface area, the position of the oxygen content and the contact resistance maxima will also vary with these factors.

The contact resistance of granular carbon oxidized in air was determined more accurately in the second conductivity cell (Fig. 2), where the four-probe method of resistance measurement effectively eliminated errors due to electrode–carbon contact resistance. The contact resistance–oxidation temperature curve (Fig. 5) obtained was similar to that obtained previously, except that the range of resistance involved was greater. The rate of decrease in resistance with temperature above 600° C was found to depend on the technique of oxidation employed. If the oxidation was carried out with the carbon in a tube sealed at one end, and change of atmosphere during cooling was prevented by a tap at the other, then lower resistance values were obtained than if the carbon were oxidized in an open boat and cooled rapidly in air by withdrawal from the furnace. The higher values in the latter method are attributed to the re-oxidation during the cooling process, when the carbon passes through the optimum temperature (530° C) for oxide formation whilst in contact with the air. This contact with the air was prevented in the closed tube, yielding lower values for the contact resistance.

A re-examination of the contact resistance–temperature curves for inert gases (Fig. 3) in terms of surface oxide

formation and decomposition indicates that the raw material undoubtedly contains a surface oxide film which begins to decompose at about 400° C and continues to do so in hydrogen from 300–800° C (Marsh 1951). The change in slope above 900° C is attributed to a structural change in the carbon itself. This decrease in resistance is primarily due to a continuation of the carbonization process used in the manufacture of the carbon granules.

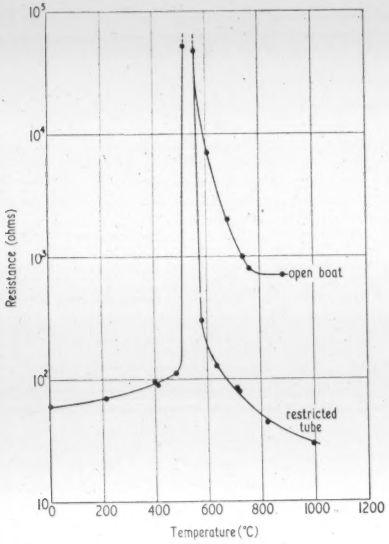


Fig. 5. Air oxidation of carbon.

The existence of an oxygen complex on the high-resistance carbon samples has been confirmed by analysis of the gases evolved from the carbon at high temperatures (900° C and above). Correlation is at present being sought between contact resistance and the surface oxygen content. It is hoped that, by suitably altering the rates of formation and decomposition of the surface oxide, the oxygen content, and hence the resistance, may be kept constant over a wide range of temperatures.

Acknowledgment

The author wishes to thank Standard Telecommunication Laboratories Ltd. for permission to publish this paper.

References

GULBRANSEN, E. A., and ANDREW, K. F., 1953, *Industr. Engng Chem.* (*Industr.*), **44**, 1048.

HIRABAYASHI, H., and TOYODA, H., 1954, *Tanso* (*Carbons*), **4**, 2.

MARSH, J. D. F., 1951, *Inst. Gas Engrs, Communication No.* 393.

PURI, B. R., SINGH, D. D., NATH, J., and SHARMA, L. R., 1958, *Industr. Engng Chem.* (*Industr.*), **50**, 1071.

SANDOR, J., 1944, *Proc. Conference on Ultra-fine Structure of Coals and Cokes, London*, p. 342.

SIEBEL, K., 1921, *Z. Phys.*, **4**, 288.

SMITH, R. N., 1959, *Quart. Rev.*, **13**, 287.

SNOW, C. N., WALLACE, D. R., LYON, L. L., and CROCKER, G. R., 1958, *Proc. 3rd Conference on Carbon* (New York: University of Buffalo), p. 279.

WAETZMANN, E., and GIGLING, O., 1928, *Akust. Z.*, **3**, 169.

WRIGHT, R. H., and MARSHALL, M. J., 1928, *Trans Amer. Electrochem. Soc.*, **54**, 149.

Influence of alloy structure on the transfer of matter

by W. MERL, Dipl.-Phys., Research Laboratory, Dr. E. Dürrwächter-DODUCO-KG., Pforzheim, Germany

Paper presented on 7th April 1961, at The Institute of Physics and The Physical Society Symposium on Electrical Contacts

MS. received 1st May 1961

Abstract

The influence of conductivity on the transfer of matter at break (6 v, 4·3 a, 0·06–20 μh) was investigated for precipitation-hardened gold alloys and a palladium alloy with super-structure. For the gold alloys the degree of transfer depends strongly on conductivity in the regions of residual transfer and of normal arc, but is hardly affected in the region of the short arc.

The importance of the alloy structure for the degree of material transfer between contacts has been emphasized for the first time in a U.S. Patent 2213397 (1940). There one can learn that by thermal treatment of a palladium–copper alloy, containing 40% wt. copper, it is possible to change the normal lattice structure with a statistic distribution of atoms into an ordered lattice. By this transformation one gains an increase of electrical and thermal conductivity by a factor of 7 (Vines and Wise 1941). Owing to the improved evolution of heat from the contact surface, the transfer of matter is reduced to approximately 2% under certain electric conditions (Keil and Merl 1957). This means in practice that the life of the contacts is prolonged by more than two magnitudes.

In the meantime, various further alloys of noble metals have been examined in this respect, such as gold–copper and gold–cobalt (Keil and Meyer 1953). In the alloy gold–copper one can observe transformations from the disordered to the ordered state similar to those in palladium–copper. In gold–cobalt the conductivity can be increased by precipitation hardening. With a suitable thermal treatment the homogeneous gold–cobalt alloy crystal transforms into gold base and cobalt base solid solutions. The additive conductivity of the latter depends on the annealing temperature, but is in any case higher than that of the homogeneous solution crystals.

Later, proof of the influence of this precipitation hardening on the transfer of matter with gold–nickel alloys containing more than 8% wt. nickel was given (Merl 1960). Fig. 1 shows the anode net gain per operation of gold–nickel contacts for a break current of 4·3 a and 3·4 μh in a 6 v circuit depending on the annealing temperature. The arc duration, measured by an oscilloscope, was 1·5 μs. The transfer of matter of the solid-solution alloy can be reduced to 10% by a thermal treatment at 500° c.

These first experiments were always carried out under similar electric conditions. However, it is clearly understood that the influence of the conductivity depends on the mechanism of the material transfer and, therefore, must change over a wide range of the arc time. Knowing these various mechanisms, one can show the influence of the conductivity by comparing the material transfer of an alloy in the two different states. One alloy from each of the systems palladium–copper, gold–cobalt and gold–nickel has, therefore, been investigated over a wide range of circuit inductance. In Fig. 2 the cathode net gain per switching is plotted against the inductance. The upper curve is for palladium–copper in the disordered state; the curve below was obtained after a thermal treatment of the contacts for

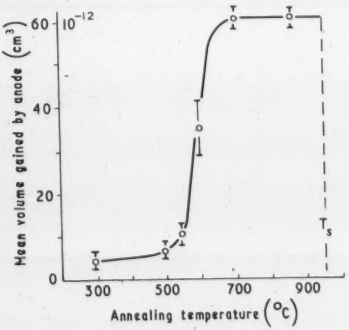


Fig. 1. Anode net gain per switching of gold–nickel 84/16 contacts depending on annealing temperature (6 v, 4·3 a, 3·4 μh, arc duration 1·5 μs).

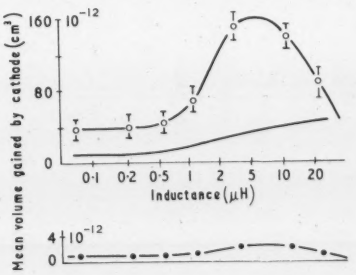


Fig. 2. Cathode net gain per switching of palladium–copper contacts (6 v, 4·3 a).

—— ○ —— Pd/Cu 60/40 disordered.
—— ● —— Pd/Cu 60/40 ordered.
———— Au.

several hours at a temperature of 500° C. The electrical conditions of the circuit again are 6 V, 4·3 A, the speed of contacts at break was in the order of 10 cm/s. Up to 1 μH the transfer of matter is independent of the inductance and is thus to be ascribed to the bridge mechanism. Between 1 and 4 μH an additional transfer of matter is caused by the short arc. At approximately 4 μH the normal arc begins and the cathode gain is reduced. The curve for the ordered state has a similar form, though the quantity is considerably smaller. For comparison the curve for gold has also been included.

In Fig. 3 the same measurements for gold–cobalt containing 3% cobalt are indicated. By precipitation hardening at

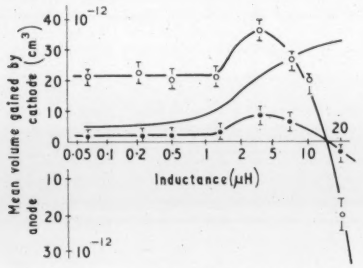


Fig. 3. Net gain per switching of gold–cobalt contacts (6 v, 4·3 A).

—— ○ —— Au/Co 97/3 homogeneous.
—— ● —— Au/Co 97/3 heterogeneous.
—— Au.

450° C the conductivity can be increased by a factor 4·5. The curves for gold–cobalt and palladium–copper are very similar. This conformity, however, is only accidental. Generally the curves of alloys with such different contents of base metal are dissimilar. The difference between the two states of gold–cobalt is not as large as that of palladium–copper, but remains appreciable.

Figure 4 shows the curves for gold–nickel with 16% wt. nickel. The transfer of matter is plotted against the arc duration for various annealing temperatures. At 300 and 500° C the curves are identical, as also the electrical conductivities. Between 500 and 700° C the transfer of matter and the conductivity depend strongly upon the annealing temperature. Annealing above 750° C and quenching makes the alloy homogeneous, and the curves again coincide. The difference in the conductivity between both the external curves is represented by a factor 2·5. In the investigated range of inductance the curves of gold–nickel and gold–cobalt are not similar. The net transfer of matter of gold–nickel from the anode to the cathode is very small, while the transfer from the cathode to the anode by the normal arc is considerable.

The point of intersection of the curve with the abscissa marks the arc duration at which the amounts of metal transported subsequently from the anode to the cathode by the short arc, and from the cathode to the anode by the normal arc, compensate. This arc duration is approximately 0·3 μs for gold–nickel and 5 μs for gold–cobalt. The corresponding values of inductance are approximately 0·4 μH for gold–nickel and 15 μH for gold–cobalt.

For the purpose of comparison it was found useful to normalize the scale of inductance. The unit was chosen such that the inductance corresponded to the point of intersection. Thus one obtains limited regions each for bridge mechanism, short arc and normal arc. It must be stressed, however, that for the same values of the normalized inductance the arc durations can differ by a factor 10.

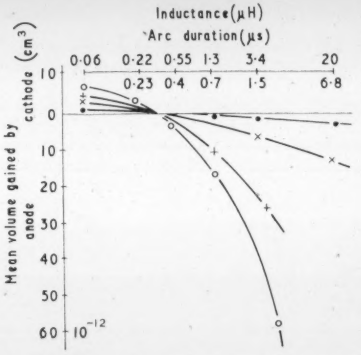


Fig. 4. Net gain per switching of gold–nickel contacts (6 v, 4·3 A).

—— ● —— Au/Ni 84/16 annealed at 300 resp. 500° C.
—— × —— Au/Ni 84/16 annealed at 550° C.
—— + —— Au/Ni 84/16 annealed at 600° C.
—— ○ —— Au/Ni 84/16 annealed at 700 resp. 850° C.

In order to find out the influence of the conductivity in the different regions, it is only necessary to consider the ratio of the volumes transferred in the two different states of an alloy. In Fig. 5 this ratio m_1/m_2 is plotted against the normalized inductance L/L_0 for the alloys gold–cobalt and

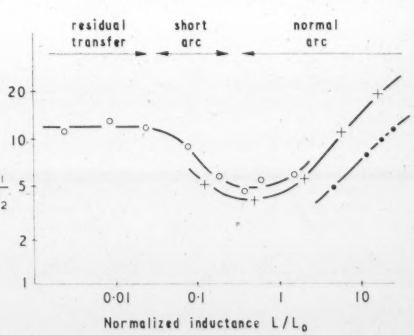


Fig. 5. Ratio m_1/m_2 of net gains for two different states of an alloy depending on the normalized inductance L/L_0.

—— ○ —— Au/Co 97/3.
—— + —— Au/Ni 84/16.
—— ● —— Au/Ni 90/10.

gold–nickel. The curves are similar near the minimum, owing to the transition from the short arc to the normal arc. At smaller values of the normalized inductance the residual transfer of gold–cobalt was measured only, while at large values the coarse transfer of two gold–nickel alloys with 16 and 10% nickel was measured. The position of the

curves depends on the increase of the conductivity by the thermal treatment, but there is no simple relation between the ratio of the conductivity and the ratio of the volumes of metal transferred.

These curves show that the conductivity of the material is of greater importance in the region of the bridge mechanism and the normal arc than in the region of the short arc. Since the break of the bridge, the short arc and the normal arc occur one after another, the region of the short arc includes the bridge transfer. If one considers the transfer by the short arc alone, the minimum would be much lower; this means that the differences in the conductivities hardly affect the amount of transfer in the region of the short arc. Investigating the material transfer at make with capacitive load, the same result is achieved. The anode arc at make is very similar to the short arc at break, except for the mechanism of ignition.

For the case of short-circuiting a capacitor, Fig. 6 shows the quantity of the material transfer versus the potential of

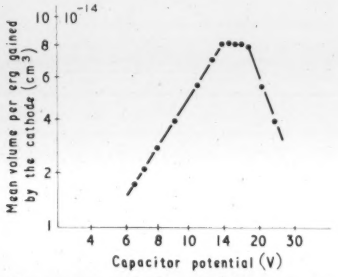


Fig. 6. Net gain per switching of gold contacts at make (1·04 μF, 0·048 μH, 58 mΩ (h.f.)).

the capacitor. The apparatus used was almost the same as described by Germer and Haworth (1949); however, the capacitance was considerably larger. Because of this the anode arc changed into a normal arc above a potential threshold, and thus the gain by the cathode was reduced. The ordinate presents the amount of metal transported from the anode to the cathode, divided by the energy of the capacitor.

The anode arc between gold contacts ignites at nearly 14 v (Merl 1956). At lower voltages the transfer of metal is due to arcs at break when microscopic peaks of both contact surfaces boil at the first touch of the metals. The second break of the curve at 17 v marks the transition from the anode arc to the normal arc. The decrease is not due to a cathode arc or an air breakdown, the voltage being too low.

Figure 7 gives the same measurements for gold–cobalt containing 5% cobalt in the homogeneous and the hetero-

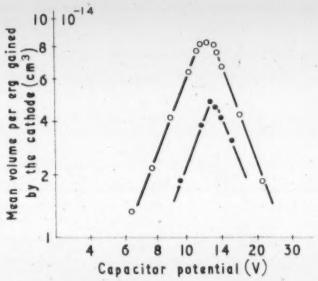


Fig. 7. Net gain per switching of gold–cobalt contacts at make (1·04 μF; 0·048 μH, 58 mΩ (h.f.)).
—— ○ —— Au/Co 97/3 homogeneous.
—— ● —— Au/Co 97/3 heterogeneous.

geneous state. The maximum of the curve for the homogeneous state is 8×10^{-14} cm^3/erg, for the heterogeneous state approximately 5×10^{-14} cm^3/erg; that is more than half the value for the homogeneous state, although the resistivity is about ten times smaller.

The experiments are not yet finished, but it is evident that the transfer of matter on break can be reduced to a great extent by increasing the conductivity of an alloy. In some cases a rise of the conductivity by a factor 3 could diminish the quantity of metal transferred to 5%. On closure, however, changes of the conductivity are of little effect on the material transfer. An increase of the conductivity by a factor 10 reduced the material transfer only by one half.

References

GERMER, L. H., and HAWORTH, F. E., 1949, *J. Appl. Phys.*, **20**, 1085.

KEIL, A., and MERL, W., 1957, *Z. Metallkunde*, **48**, 16.

KEIL, A., and MEYER, C.-L., 1953, *Z. Metallkunde*, **44**, 22.

MERL, W., 1956, *Z. Naturforsch.*, **11a**, 1041.

MERL, W., quoted in KEIL, A., 1960, *Werkstoffe für elektrische Kontakte* (Berlin/Göttingen/Heidelberg: Springer-Verlag), p. 253.

VINES, R. F., and WISE, E. M., 1941, *The Platinum Metals and their Alloys* (New York: International Nickel Company), p. 108.

Measurement of the dielectric constant of filament nylon in a transverse field

by F. S. WARD,* M.A.Cantab., A.Inst.P., The British Cotton Industry Research Association, Shirley Institute, Manchester

Abstract

An instrument has been designed for measuring the dielectric constant of a material in the form of a single cylindrical filament, and experiments are described in which it is used for measurements on nylon. The filament is placed centrally between two silvered optical flats, which form a parallel-plate capacitor and serve as the tuning capacitor of a radio-frequency oscillator. According to an existing theory, the change in capacity due to the presence of the dielectric is given by:

$$\delta C = \left(\frac{b^2 H}{2D^2}\tanh\frac{\pi B}{2D}\right)\Big/\left(\frac{\varepsilon+1}{\varepsilon-1}-\frac{\pi^2 b^2}{3D^2}\right)$$

(where b is the radius of the filament, H the length of the capacitor, B the width and D the distance between the plates) and δC can be determined from the corresponding change in oscillator frequency, if the electrical parameters of the circuit are known. The experimental observations confirmed the general form of the equation, but they were not sufficiently accurate to verify the second-order term $\pi^2 b^2/3D^2$. A value of 3·0 was found for the dielectric constant of nylon at 6 to 7 Mc/s, which is in reasonable agreement with values of 3·1 to 3·4 recorded by other workers for both filament and bulk nylon.

Introduction

It is of interest, on both theoretical and practical grounds, to know the dielectric constant of a material such as nylon which exists in bulk and in filament form. Measurements have been made of the dielectric constant of nylon in the form of disks or rods (Rushton and Russel 1956), but the only data available for nylon filaments are derived from measurements on assemblies of fibres. The difficulty with this type of experiment is that it gives an overall dielectric constant of the fibres in conjunction with the surrounding medium, and the true dielectric constant of the fibres can only be determined if the field distribution throughout the mixture is known. Rayleigh (1892) has solved the problem for an infinite rectangular array of dielectric cylinders, but his result does not seem to have been used in connection with fibres.

Errera and Sack (1943) avoided the difficulty by surrounding the fibres with a liquid whose dielectric constant was the same as that of the fibres themselves, but their results are open to criticism on the grounds that the nylon may have been affected by the surrounding medium. In some of the experiments made at 3000 Mc/s by Shaw and Windle (1950), the nylon filaments were arranged, in a resonant cavity, with their axes parallel to the electric field and these results lead directly to a value for the dielectric constant of nylon at this microwave frequency; but in other measurements, made with the electric field perpendicular to the fibre axis, they found it necessary to use an empirical equation in order to derive the true dielectric constant. The measurements made by Hearle (1954, 1956) at low and radio frequencies also give only an 'apparent dielectric constant' of nylon.

It appeared useful therefore to supplement the existing data by measuring the dielectric constant of nylon filaments in a transverse electric field at frequencies well below the microwave region, and recent theoretical calculations by Mack (1955a, b) made this possible. Although the method now described was only used for nylon, it also applies to any fibre with a circular cross section.

Apparatus

The method of measurement is to bring the filament whose dielectric constant ε is to be determined into the centre of a parallel-plate capacitor and measure the resulting increase δC_x in capacity which is given by the equation

$$\delta C_x = \left(\frac{b^2 H}{2D^2}\tanh\frac{\pi B}{2D}\right)\Big/\left(\frac{\varepsilon+1}{\varepsilon-1}-\frac{\pi^2 b^2}{3D^2}\right) \qquad (1)$$

where b is the radius of the filament, H the length of the capacitor, B the width and D the distance between the plates. Although Eqn (1) applies strictly to a section of a capacitor with infinite plates, it is likely to give a very close approximation to the change in capacity of the capacitor used in this experiment (see Appendix 3). The increase in capacity is inevitably small because of the fineness of a single filament and if it is to be measurable, the capacitor must be long and narrow, with the plates set very close together (e.g. 0·007 cm to 0·014 cm). With such a small separation, the plates must be accurately parallel and as flat as possible. This was achieved by using silvered optical flats as the capacitor plates, and an interferometer device to set them parallel.

To ensure adequate conductivity of the capacitor plates, the glass surfaces were silvered much more heavily than for normal optical purposes, and the necessary low-resistance contacts with the silver were produced by extending the coating to cover 'terminals' of silver wire set in the bevels at the top and bottom of the plates. By masking the glass surface suitably with two straight edges during the silver evaporation, the 'live' capacitor plate was restricted to a strip one millimetre wide, close to one edge of the first glass plate. A much broader strip was deposited on the opposing surface of the other optical flat to serve as the earthed plate of the capacitor.

* Now at Ferranti Ltd., Wythenshawe, Manchester 22.

The instrument in which the capacitor plates are mounted is based on a Hilger design (Pollard 1929) and uses kinematic principles, but its form was modified considerably to meet the particular requirements of this experiment. A drawing of the apparatus can be seen in Fig. 1. Each glass plate A, B is supported by a horizontal table C, D respectively; one of these C can be tilted about two axes at right angles, and the other D translated in a direction normal to the plates in order to set the two optically flat surfaces accurately parallel with any desired spacing. The control screws E, F, G which are used for these adjustments are provided with micrometer-type screw-threads for fineness of control, and a further ten to one reduction in movement is produced by first-order levers H, J, K respectively.

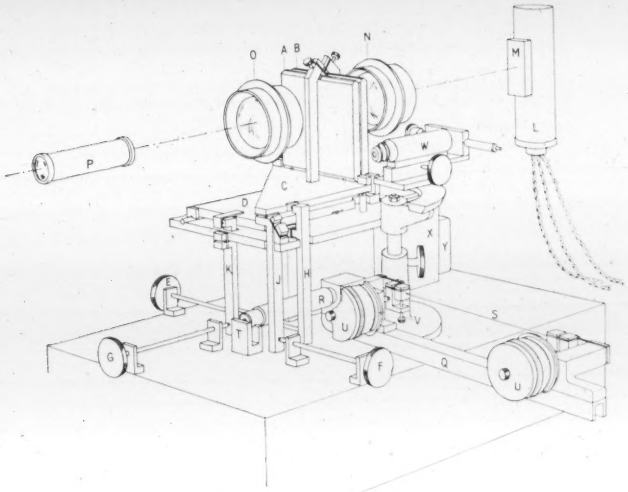


Fig. 1.

Any lack of parallelism of the plates can be detected by noting the fringe pattern, which is formed by multiple reflections at the optical flats. These were partially silvered in the areas not already occupied by the capacitor plates in order to make the fringes sharper and more intense. The optical system used is shown schematically in Fig. 1. A sodium lamp L, with a diffusing screen M, provides the illumination and two 14 in. f/5·6 lenses act as collimator N and collecting lens O. The fringes are viewed by transmitted light. If the air-space between the optical flats is wedge-shaped, the 'contour' fringes which are formed in or near the plates can be seen directly through the sighting tube P, but when the two surfaces are exactly parallel, a new system of circular ('Haidinger') fringes is formed at infinity, and to see these it is necessary to use additional lenses as shown in the drawing. These are two simple lenses, the objective having a focal length of 10 cm and the eye-piece 20 cm.

With the smallest values of the plate separation D used in this experiment, the position of the nylon filament in the capacitor is very critical and the maximum deviation from the mid-plane which can be tolerated is about 0·01 mm (Appendix 1). The mechanism which is used to transport the filament into and out of the capacitor was kinematically designed and operated well within the prescribed tolerance. It consists of an arm Q, clamped at one end to a horizontal shaft R which rests in two vee-blocks T and can therefore rotate, but is located in the axial direction by associated pressure plates. The specimen S is held between two mounting stages which are attached to this arm by spring-steel hinges and each end of the filament can be moved independently in two directions by adjusting screws. Micrometer heads U are used for the more critical adjustment of bringing the filament into the mid-plane between the capacitor plates, and set-screws V for movement parallel to the plates.

With the microscope W it is possible to observe the nylon filament while it is actually in the capacitor and with suitable illumination, multiple images of the filament can be seen, due to successive reflections at the silvering, and these provide a sensitive indication of its centrality between the plates. The microscope can be raised by a rack-and-pinion to cover the whole extent of the capacitor, and by sliding the micro-

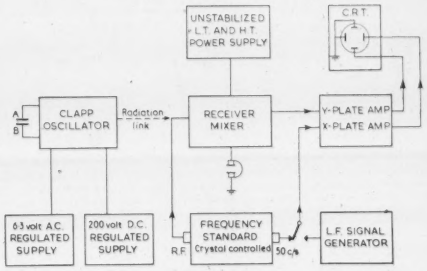


Fig. 2.

scope carriage over a quadrant its axis can be turned through a right angle, and the filament viewed against the background of the 'live' capacitor plate. For this purpose the other plate B can be displaced slightly to one side. A micrometer eyepiece is fitted to the microscope for measurement purposes.

Figure 2 shows the general arrangement of the apparatus in schematic form. Since the capacity increase (δC_x), which has to be measured, is so small (0·05 pF), a resonant frequency method was adopted in which the silvered glass plates A and B form the tuning capacitor of a radio-frequency oscillator (C_1 in Fig. 3) and a change in capacity is transformed

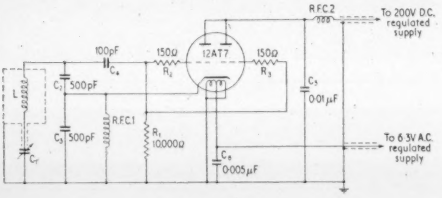


Fig. 3.

into a change in frequency, which can be measured accurately. The circuit of the Clapp oscillator was chosen because of its low tuning capacity and good frequency stability. If C is the effective capacity in parallel with a coil, of inductance L

$$4\pi^2 f^2 LC = 1 \tag{2}$$

and hence

$$-\frac{\delta f}{f} = \frac{\delta C}{2C} \tag{3}$$

where f is the frequency of oscillation. Since C includes the stray capacity of the circuit this should be kept as small as possible for maximum sensitivity. Accordingly a rigid, air-spaced coaxial lead of low capacity is used to connect the capacitor to the coil which is screened by a copper box of suitable dimensions. The other components of the oscillator are suitably mounted on a separate chassis in an arrangement which minimizes the effect of the valve heat-dissipation on the rest of the circuit. Operating conditions are kept as nearly constant as possible by the use of stabilized power supplies. By choosing suitable values for the space between the capacitor plates, the oscillator frequency can cover the range of 6 to 7 Mc/s.

In order to increase the accuracy of measurement, the oscillator signal is mixed with an approximately equal standard frequency to produce a difference frequency which is in the audible range. The frequency standard which was used for this purpose is based on a 100 kc/s crystal, and its r.f. output contains all the harmonics of this frequency. The measurement procedure is therefore to tune the receiver to the desired frequency, which can be any multiple of 100 kc/s in the range 6 to 7 Mc/s and adjust the Clapp oscillator until the beat frequency falls to zero. When the nylon sample is brought into the centre of the measuring capacitor, the resulting change δf in oscillator frequency appears directly as a beat frequency. It can be heard in the headphones and measured, by the method of Lissajous figures, with the oscillograph which forms part of the frequency standard, and a low-frequency signal generator. The scale of the signal generator was checked against a 1 kc/s signal from the frequency standard, by the same method.

The receiver scale too was calibrated against the frequency standard with an auxiliary 3·5 Mc/s crystal oscillator to produce marker signals; in an actual measurement of dielectric constant, however, the receiver scale is only used to identify the particular harmonic of 100 kc/s which is being used.

When adjusting the Clapp oscillator for zero beat frequency, it was found convenient to have a visual indicator as well as the headphones, and for this purpose a 50 c/s signal was applied to the X-plates of the oscilloscope.

As shown in Appendix 2, the circuit of the Clapp oscillator is equivalent to Fig. 4, in which C_x represents the measuring capacitor (formed by the silvered glass plates), and C_0 its stray capacity. $C_x + C_0$ is therefore identical with C_1 of Figs 3 and 7(*a*). C_L is the stray capacity across the coil, and C_A is the effective capacity of the two series capacitors C_2, C_3, together with the capacitors C_4, C_5 and the associated

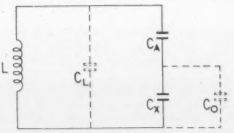


Fig. 4.

inter-electrode capacities of the valve (see Fig. 7(*a*)). Hence the equivalent tuning capacity C, across the coil is given by the equation:

$$C = C_L + \frac{C_x + C_0}{1 + (C_x + C_0)/C_A} \tag{4}$$

and hence

$$dC = \frac{dC_x}{\{1 + (C_x + C_0)/C_A\}^2}.$$

Combining this result with Eqns (1) and (3), and putting $\tanh \pi B/2D = 1$ (see Appendix 1)

$$-\frac{f}{C\{1 + (C_x + C_0)/C_A\}^2 \delta f} = \frac{4D^2}{b^2 H}\left(\frac{\varepsilon + 1}{\varepsilon - 1}\right) - \frac{4\pi^2}{3H}. \tag{5}$$

Measurement of circuit constants L, C_L, C_A

In order to find the circuit constants of the Clapp oscillator, the measuring condenser C_x was replaced by a variable air condenser C_x' and the relation between its capacity and the oscillator frequency f was determined over a range of frequencies which included those actually used in the experiment and extended some distance on either side. Although the graph of $1/f$ against C_x' was not linear, due to the effect of the coupling condensers, it was possible to extrapolate the curve slightly and find the stray capacity C_0' associated with the leads to the calibration capacitor. According to Eqns (2) and (4), the relation between $1/f^2$ and

$$\frac{C_x' + C_0'}{1 + (C_x' + C_0')/C_A}$$

should be linear and this was found to be the case provided that the calculated value of 270 pF was used for C_A (see Appendix 2). From the gradient and intercept of the straight line, the inductance L of the coil was found to be $18{\cdot}2 \times 10^{-6}$ H, and its stray capacity C_L, 6·3 pF. With larger coupling capacitors ($C_2 = C_3 = 980$ pF) a linear relationship was again found, and values of $18{\cdot}5 \times 10^{-6}$ H and 6·55 pF were found for L and C_L respectively.

Although not required for the determination of the dielectric constant, it was interesting to calculate C_0, the stray

capacity associated with the tuning capacitor. It proved to be rather larger than expected but was sensibly constant, ranging from 14·7 cm at 6·0 Mc/s to 14·9 cm at 6·5 Mc/s and 14·6 cm at 7·0 Mc/s.

The circuit constants C_L, C_A are probably less accurately known than the inductance L, but they have less effect on the final determination of dielectric constant (see Eqn 5).

Measurement of linear dimensions *b*, *H*, *B*, *D*

The radius b of each filament was determined by weighing, assuming a density of 1·14 g cm^{-3} for nylon. Two filaments, A and B, from the same multifilament yarn, were each subdivided into shorter lengths (A1, A2, B1, B2) for the dielectric measurement. Neither filament seemed to vary significantly in diameter along its length, but B was about 10% larger than A. Their radii were 9·83 (A1), 9·89 (A2), 10·87 (B1) and 10·86 (B2) microns respectively.

The dimensions of the capacitor plates were measured with a travelling microscope and the results are: $B = 0{\cdot}1082$ cm and $H = 11{\cdot}82$ cm for the live plate and $H = 11{\cdot}76$ cm for the earthed plate. The distance D between the plates is variable and must be determined for each frequency at which dielectric measurements are made. Changes in D can be measured accurately, using the interferometer system, by counting the number of fringes which appear as the plates are moved, and this method was used to determine increases in D as the oscillator frequency was raised in successive steps of 100 kc/s. The actual values of D in arbitrary units can be determined with the microscope W and its micrometer 'eyepiece'. It was not possible to focus on the edge of the plates, but a filament was put between the capacitor plates so that measurements could be made of the distance ($2D$) between images of the filament. The magnification of the microscope had been determined previously, but it was found to vary so much with the setting of the focusing control that no reliance could be placed on the calibration. By measuring the distances between images at two different frequencies, however, without refocusing the microscope in between, the ratio D_2/D_1 of the two distances could be determined, and from this their absolute values calculated using the results for $D_2 - D_1$ already obtained.

The resulting values for D at the various frequencies are shown in Table 1; the two sets of figures correspond to two separate determinations of the ratio D_2/D_1, the first at 6·6 and 6·1 Mc/s ($D_2/D_1 = 1{\cdot}724$), and the second at 6·7 and 6·4 Mc/s ($D_2/D_1 = 1{\cdot}428$).

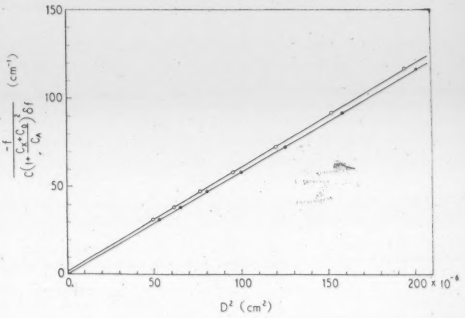


Fig. 5.

and D^2, and this is confirmed by the results, which were obtained with sample B1 (see Fig. 5). The lines, corresponding to the two sets of readings for D given in Table 1, have slightly different gradients and the derived figures for the dielectric constant are 2·98 and 2·91 giving a mean value of 2·95. Neither line intersects the y axis at the point $-4\pi^2/3H$ (i.e. $-1{\cdot}1$) predicted by the theory, however, but in view of the extent to which errors in D affect the intercept, this discrepancy is clearly not significant.

The theory also predicts that $\delta f \propto b^2$ (neglecting the small term $4\pi^2/3H$) and this, too, is borne out by the results, within the limits of experimental error. Thus, taking the average of the measurements for A1 and A2, $\delta f/b^2 = 33{\cdot}8 \times 10^8$ cm^{-2} s^{-1}, while the corresponding value for B1 is $33{\cdot}9 \times 10^8$ cm^{-2} s^{-1}.

Discussion

The results show a satisfactory agreement between theory and experiment as far as first-order terms are concerned, even though the capacitor plates are not infinite as the theory assumes. Experimental errors are too large to allow the second-order term in Eqn (1) to be checked.

The value of 3·0 for the nylon filaments is slightly less than the figures given by Rushton and Russel (1956) for bulk nylon (3·1 to 3·4 according to the moisture content), by Errera and Sack (1943) (3·15 to 3·18) for filaments in a transverse electric field and by Shaw and Windle (1950) (3·09 to 3·18) for an axial field. In view of the cumulative effect of experimental errors on the value of the dielectric constant derived from the present experiment, however, these differences are not significant.

Table 1. *Distance D between capacitor plates at different frequencies f*

Frequency f(Mc/s)	6·0	6·1	6·2	6·3	6·4	6·5	6·6	6·7	6·8	6·9	7·0
D (cm)	0·00660	0·00730	0·00809	0·00898	0·01001	0·01120	0·01258	0·01419	0·01608	0·01837	0·02115
	0·00636	0·00706	0·00785	0·00875	0·00977	0·01096	0·01234	0·01395	0·01584	0·01813	0·02091

Measurement of dielectric constant

The dielectric constant of a filament can be calculated from the foregoing parameters and the change δf in oscillator frequency on bringing the filament into the measuring capacitor (see Eqn 5). This ranged from 1·72 kc/s (when $f = 6{\cdot}70$ Mc/s) to 4·61 kc/s (when $f = 6{\cdot}10$ Mc/s).

According to Eqn (5) there should be a linear relation between

$$\frac{-f}{C\{1 + (C_x + C_0)/C_A\}^2 \delta f}$$

Appendix 1

The change in capacity of a parallel-plate capacitor due to a non-central cylindrical dielectric

The change δC_x in the capacity of a parallel-plate capacitor of length H due to the presence of a dielectric cylinder of

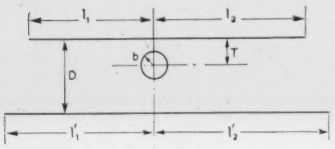


Fig. 6.

radius b between the plates in the position shown in Fig. 6 is given by the equation:

$$\delta C_x = \frac{b^2 H}{8D^2} \cdot \frac{\dfrac{\sinh(\pi l_1/D)}{\cosh(\pi l_1/D) - \cos(\pi T/D)} + \dfrac{\sinh(\pi l_2/D)}{\cosh(\pi l_2/D) - \cos(\pi T/D)} + \dfrac{\sinh(\pi l'_1/D)}{\cosh(\pi l'_1/D) + \cos(\pi T/D)} + \dfrac{\sinh(\pi l'_2/D)}{\cosh(\pi l'_2/D) + \cos(\pi T/D)}}{(\varepsilon+1)/(\varepsilon-1) - (\pi b/2D)^2\{\mathrm{cosec}^2(\pi T/D) + \tfrac{1}{3}\}} \qquad (6)$$

neglecting terms of the order of b^6. (This is equivalent to Eqn 2.2 in the paper by Mack (1955b).) Since the maximum value of a cosine is 1, and cosh $(1{\cdot}7\,\pi) = 100$, variations in T cannot change the numerator by more than 1% if $l/D \geqslant 1{\cdot}7$, and with the values of 2·5 or more adopted in this experiment, the effect is less than 0·1% and can be ignored. The denominator can be written as

$$\frac{\varepsilon+1}{\varepsilon-1} - \frac{\pi^2 b^2}{3D^2} - P^2 \text{ where } P = \frac{\pi b}{2D}\cot\left(\frac{\pi T}{D}\right)$$

and hence if the effect on δC_x of deviations from the mid-plane is not to exceed 0·5% then P must not be more than 0·1 (assuming $\varepsilon = 3$). Thus, if $b = 0{\cdot}001$ cm, and $D = 0{\cdot}0073$ cm (the smallest spacing used for the actual measurement of dielectric constant), T/D must not be less than 0·362. This corresponds to a deviation $\delta(= \frac{1}{2}D - T)$ from the mid-plane of $0{\cdot}138D$ or 0·001 cm. The maximum permissible deviation when $D = 0{\cdot}014$ cm (the largest spacing used in the dielectric measurement) is 0·0033 cm.*

The effect on δC_x of displacing the filament parallel to the plates is less marked. For example, putting $T/D = 1/2$ in Eqn (6) we obtained the simpler relation

$$\delta C_x = \frac{b^2 H}{8D^2} \cdot \frac{\Sigma \tanh(\pi l/D)}{(\varepsilon+1)/(\varepsilon-1) - \pi^2 b^2/3D^2}$$

and since tanh $(0{\cdot}955\,\pi) = 0{\cdot}995$, tanh $(\pi l/D)$ will differ from 1 by not more than 0·5% if $l \geqslant 0{\cdot}955D$. Even in the least favourable case, therefore, where $D = 0{\cdot}014$ cm and $l_1 + l_2 \simeq 7D$ the deviation from the centre of the capacitor can be as much as $2{\cdot}5D$ (i.e. 0·35 mm) without introducing an error of more than 0·5% in δC_x.

Appendix 2

The circuit of the Clapp oscillator (Fig. 3) is equivalent to Fig. 7(*a*) where C_{gk}, C_{ak}, C_{ag} and C_{hk} are, respectively, the grid–cathode, anode–cathode, anode–grid, and heater–cathode capacities of the valve. (The effect of the choke R.F.C.1 is ignored since its resonant frequency is very much higher than the oscillator frequency; R_1, R_2 and R_3 are also neglected in considering the equivalent capacity.)

The portion of the circuit (shown dotted) between A and B is equivalent to Fig. 7(*b*), where $C_p = C_{gk} + C_{ag}(1 + C_{gk}/C_{ak})$ and $C_Q = C_{ak} + C_{ag}(1 + C_{ak}/C_{gk})$. This result is obtained by replacing the capacity C_{ag} by the equivalent pairs of capacitors, in series, whose capacities are in the same ratio as C_{gk} and C_{ak}, and combining them with C_{gk}, C_{ak} respectively. There is, therefore, an effective capacity $C_4 C_p/(C_4 + C_p)$ in parallel with C_2, and a capcity $C_5 C_Q/(C_5 + C_Q)$ in parallel with C_3 and C_{hk}. Hence the circuit to the right of DE in Fig. 7(*a*) can be replaced by a single capacity C_A (as in Fig. 4), where:

$$\frac{1}{C_A} = \frac{1}{C_2 + C_4 C_p/(C_4 + C_p)} + \frac{1}{C_3 + C_5 C_Q/(C_5 + C_Q) + C_{hk}}. \qquad (7)$$

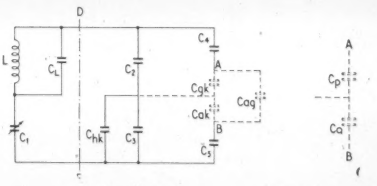


Fig. 7.

With the measured values: $C_2 = C_3 = 496$ pF, $C_4 = 100$ pF, $C_5 = 0{\cdot}01\ \mu$F and assuming the figures given by the valve data sheet: $C_{ag} = 3{\cdot}4$ pF, $C_{ak} = 0{\cdot}36$ pF, $C_{hk} = 4{\cdot}8$ pF and $C_{gk} = 67{\cdot}5$ pF (which allows for the Miller Effect, assuming a stage-gain of 25 for the two sections of the valve combined), Eqn (7) gives $C_A = 270{\cdot}6$ pF. Similarly if $C_2 = C_3 = 980$ pF and the other parameters are unchanged, $C_A = 513{\cdot}4$ pF. This compares with the value of 508 pF which was found empirically to give the best linear relation between $1/f^2$ and $(C'_x + C'_0)/\{1 + (C'_x + C'_0)/C_A\}$.

Appendix 3

(a) Edge-effects in parallel-plate capacitors

From the theory of edge effects in parallel-plate capacitors (Jeans 1920) the field E at the point $(x, 0)$ in the mid-plane between two semi-infinite conducting planes is given by the equation:

$$E = \frac{\pi R/h}{e^{u/R} + 1}$$

where

$$x = -\frac{h}{\pi}\left(1 + e^{u/R} + \frac{u}{R}\right) \qquad (8)$$

* Substantially the same information is available from Table II(*a*) of the Memoir by Mack (1955b) which, allowing for slight differences in nomenclature, tabulates

$$\left\{1 - \frac{\varepsilon-1}{\varepsilon+1}\left(\frac{\pi^2 b^2}{3D^2} + P^2\right)\right\}^{-1} - 1$$

as fractions of a thousand.

Mack agrees that an error has arisen in the equation given in section 3.1 of his Memoir which should read:

$$\xi = 2D^2\delta C/kb^2 = 8D^2 A/kb^2.$$

Hence, if we equate the numerator in Eqn (6) to four, and put $\delta C_x/H = \delta C$, we find that $\xi - 1$ is equal to the expression given above.

and the potential difference between the plates is $2\pi R$ (see Fig. 8). As $u \to -\infty$, $x \to +\infty$ and $E \to \pi R/h (= E_\infty$ say). From the above equations we have, when $x/h = 1{\cdot}14$, $E/E_\infty = 0{\cdot}990$, and when $x/h = 1{\cdot}88$, $E/E_\infty = 0{\cdot}999$.

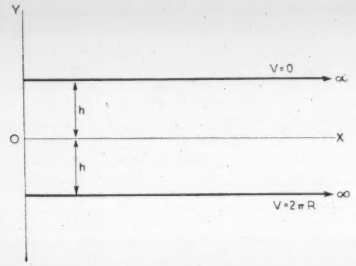


Fig. 8.

Hence the field mid-way between the plates is sufficiently uniform at distances greater than about $2h$ from the edge, where $2h$ is the separation of the plates.

Similarly, the field mid-way between an infinite and a semi-infinite plane reaches 0·9995 of its maximum value at a distance from the edge equal to the separation of the plates.

of the parallel-plate capacitor due to the edge. There are two cases: (i) where one plate is semi-infinite and the other infinite, i.e. $D = H$ and $V = \pi R$; (ii) where both plates are semi-infinite, i.e. $D = 2h$ and $V = 2\pi R$.

Table 2 gives both these corrections (Δ_1, Δ_2 respectively) for a strip of the same width (B) as the capacitor used in this experiment, and with values of D appropriate to the given frequencies. The outer edge of the measuring capacitor, where both plates terminate simultaneously, is assumed to correspond to case (ii), and the inner edge, where the earthed plate continues well beyond the other, to case (i). As a first approximation, the capacity C_x of the capacitor is taken to be $(HB/4\pi D) + \Delta_1 + \Delta_2$.

(b) Edge-effects in a parallel-plate capacitor terminating in 45° bevels

In this case the field E at a point $(x, 0) = \pi R/rh$, where $x = -4r + \log(r+1)/(r-1) + 2\tan^{-1}r$, and as r goes from ∞ to 1, x increases from $-\infty$ to $+\infty$ and E from 0 to $\pi R/h$. From these equations $E/E_\infty = 0{\cdot}999$ when $x/2h = 0{\cdot}82$, and hence the field is substantially uniform beyond this point. By considering the charge distribution on the plates, and inserting values for the dimensions of the capacitor which are appropriate to this experiment, the increase in capacity due to the bevels can be shown to vary from $0{\cdot}19B$ to $0{\cdot}26B$ over the range of dimensions D given in Table 2. Since $B = 0{\cdot}1$ cm, this correction is small.

Table 2

f (Mc/s)	6·0	6·2	6·4	6·6	6·8	7·0
D (cm)	0·00660	0·00809	0·01001	0·01258	0·01608	0·02115
$HB/4\pi D$ (cm)	15·41	12·57	10·16	8·08	6·33	4·81
Δ_1 (cm)	1·50	1·45	1·39	1·33	1·26	1·19
Δ_2 (cm)	0·85	0·82	0·79	0·76	0·72	0·69
C_x (cm)	17·76	14·84	12·34	10·17	8·31	6·69

In the present experiment, therefore, with the limits on centring specified in Appendix 1, the nylon filament is brought into a region of uniform electric field, and the dipole moment induced in the filament will be the same as if the capacitor plates were infinite. According to Mack, the charges induced in the plates by the dipole are very localized, and the resulting increase in capacity of the capacitor is likely to be very close to the value given by Eqn (1).

The capacity of the capacitor shown in Fig. 8 can be calculated from the equation:

$$q = H(u_1 - u_2)/4\pi \qquad (9)$$

where q is the charge on the portion of the conductor bounded by the lines $u = u_1$, $u = u_2$ and H is its length in the direction perpendicular to the xy plane. For the upper plate:

$$x = \frac{h}{\pi}\left(e^{u/R} - \frac{u}{R} - 1\right) \qquad (10)$$

and if u_1, u_2 are the solutions of this equation for a given value of x, then q is the charge on the strip of width x bordering the edge. The corresponding capacity is q/V, where V is the potential difference between the plates and hence if $\Delta = (q/V) - (Hx/4\pi D)$, Δ represents the increase in capacity

According to Mack, the result given in Eqn (1) still applies, when the dielectric cylinder is brought into this type of capacitor, but the effective length H of the capacitor is reduced by $0{\cdot}075B$. This is negligible in the present experiment.

References

Clapp, C. W., 1948, *Proc. Instn Radio Engrs, N.Y.*, **36**, 3020.

Errera, J., and Sack, H. S., 1943, *Industr. Engng Chem.*, **35**, 712.

Hearle, J. W. S., 1954, *J. Text. Res.*, **24**, 307.

—— 1956, *J. Text. Res.*, **26**, 108.

Jeans, J. H., 1920, *The Mathematical Theory of Electricity and Magnetism*, 4th edn (Cambridge: University Press), p. 274.

Mack, C., 1955a, *Brit. J. Appl. Phys.*, **6**, 59.

—— 1955b, *Shirley Institute Memoirs*, **28**, 91.

Pollard, A. F. C., 1929, *The Kinematical Design of Couplings in Instrument Mechanisms* (London: Adam Hilger), p. 30.

Rayleigh, J. W., 1892, *Phil. Mag.*, **34**, 481.

Rushton, E., and Russel, G., 1956, *Electrical Research Association Technical Report* L/T 355.

Shaw, T. M., and Windle, J. J., 1950, *J. Appl. Phys.*, **21**, 956.

Spectroturbidimetry of emulsions

by J. D. S. GOULDEN, B.Sc., Ph.D., National Institute for Research in Dairying, University of Reading

MS. received 10*th May* 1961, *in revised form* 6*th June* 1961

Abstract

A theoretical treatment of the turbidity of monodisperse emulsions is discussed and its application to polydisperse systems indicated. Modifications required to commercial spectrophotometers for carrying out spectroturbidimetric studies are described. The determination of both mean globule size and concentration of the suspended phase in homogenized emulsions is illustrated.

1. Introduction

Recent advances in the theory of light-scattering, together with the increased availability of commercial spectrophotometers, have led to a renewal of interest in the spectroturbidimetry of emulsions. Although light-scattering theory is only readily applicable to emulsions in which the suspended particles are all of the same size, many emulsions can be treated as monodisperse suspensions from the point of view of light-scattering. Most of the original studies were carried out using suspensions of polystyrene spheres of uniform particle sizes, but more recent work has demonstrated the value of the method in the study and quantitative analysis of homogenized milk and other dairy products.

An emulsion owes its characteristic 'milky' appearance to the effects of light-scattering, which cause some of the incident radiation to be deviated from its original direction. Such scattering phenomena are always present when a discontinuity in refractive index occurs, but if the refractive indices of the two phases are equal, scattering is eliminated and the emulsion appears as a clear liquid.

In the case of coloured emulsions, either the suspended phase or the suspending medium shows absorption in the visible region of the spectrum. The theoretical treatment of scattering in the presence of absorption by the suspended phase is very complex, and it is fortunate that many emulsions of biological or technical interest show little or no visible absorption. When absorption bands do occur, it is often possible to avoid their effects by selecting regions of the spectrum where absorption is absent.

2. Theory of light-scattering by emulsions

In applying light-scattering theory to emulsions, it is assumed that the suspended particles are spherical in shape and that they are sufficiently far apart to keep errors due to multiple and coherent scattering below the photometric errors. This can readily be ensured experimentally by selecting emulsion dilutions for which the optical density is directly proportional to the concentration of the suspended phase.

Lothian and Chappel (1951) have shown that the attenuation of a beam of radiation by a suspension of non-absorbing spheres of radius r at any particular wavelength is given by

$$I = I_0 \exp(-\pi r^2 l K \nu) \tag{1}$$

where I_0 is the intensity of the incident beam, I is the intensity of the transmitted beam, l is the optical path length and ν is the number of spheres per unit volume. The factor K is the scattering coefficient, defined as the ratio of the scattering cross section to the geometrical cross section and has values between 0 and about 5. The optical density D is defined as

$$D = \log(I_0/I) \tag{2}$$

so that

$$D = \pi r^2 l K \nu \log e. \tag{3}$$

If the concentration of the suspended phase in the original emulsion is x g/100g emulsion and if each gram of emulsion is diluted to a total volume of y ml. before spectroturbidimetric analysis

$$(4/3)\pi r^3 \nu \rho = 0{\cdot}01(x/y) \tag{4}$$

where ρ is the density of the suspended phase. Elimination of ν from Eqns (3) and (4) gives

$$D = 65{\cdot}15(x/y)(l/\rho)(K/d) \tag{5}$$

where l, the optical path length, is expressed in centimetres and d, the globule diameter, is expressed in microns. It is often convenient to express the turbidity of a diluted emulsion in terms of $E^{1\%}_{1\,\mathrm{cm}}$, i.e. the optical density of a 1 cm path length of a diluted emulsion containing 1% (wt./vol.) of suspended phase,

$$E^{1\%}_{1\,\mathrm{cm}} = \frac{D}{l}\left(\frac{y}{x}\right). \tag{6}$$

Substitution of D from Eqn (5) gives

$$E^{1\%}_{1\,\mathrm{cm}} = 65{\cdot}15\left(\frac{1}{\rho}\right)\left(\frac{K}{d}\right). \tag{7}$$

In the particular case of homogenized milk, it is convenient to carry out the dilution by volume and to obtain an $E^{1\%}_{1\,\mathrm{cm}}$ value which includes a factor for the density of milk. This expression has previously been called the turbidity index and given the symbol T (Goulden 1960) to avoid confusion with the more general case of $E^{1\%}_{1\,\mathrm{cm}}$. Some other workers prefer to define the symbol τ for the specific turbidity from the equation

$$I = I_0 \, \mathrm{e}^{-\tau l} \tag{8}$$

i.e. τ equals $\pi r^2 K \nu$ of Eqn (1).
Hence it can be shown that

$$\tau = 2{\cdot}303\left(\frac{D}{l}\right) = 2{\cdot}303\left(\frac{x}{y}\right)E^{1\%}_{1\,\mathrm{cm}}. \tag{9}$$

Theoretical calculation of K values

As early as 1908, Mie derived exact theoretical equations for the evaluation of K values for colloidal spheres. At the time, these equations were too complex to be evaluated for all except the smallest d values and it is only since the introduction of electronic computers that accurate K values for the larger particles have become available. Two sets of accurate K values have been published. Those of Heller (1957a,b), Heller and Pangonis (1957) and Heller and McCarthy (1958) are particularly useful for emulsion studies and the data of Gumprecht and Sliepcevich (1951) apply to particles of higher refractive indices.

A most useful approximate solution has been derived by van de Hulst (1957) from the principles of physical optics. This solution applies particularly to cases where the globule sizes are of the same order as, or greater than, the wavelength of the radiation and where m, the ratio of the refractive index of the suspended phase (n_s) to that of the suspending medium (n_m) approaches unity. For many homogenized emulsions, the globule diameters are in the range 0·5 to 3 microns and since $m \simeq 1{\cdot}1$, the van de Hulst equation applies with a fair degree of accuracy for radiation in the visible region. In practice, the exact Heller–Mie solution is normally used to the limit of the published data, i.e. $d \simeq 2\ \mu$, and the van de Hulst equation for the larger globule diameters.

The factor K is a function of both m and the globule-to-wavelength size ratio, usually expressed by α defined as

$$\alpha = \frac{\pi d}{\lambda_m} = \frac{\pi d n_m}{\lambda_0}$$

where λ_m is the wavelength of the radiation in the medium of refractive index n_m and λ_0 is the wavelength *in vacuo*. For practical purposes, λ_0 can be taken as the wavelength in air, i.e. as marked on the spectrophotometer wavelength scale.

In general, K is an oscillatory function of α approaching 2 at very high α values. Goulden (1958b) has used the van de Hulst equation to evaluate the (K, α) curve for milk-fat in water where $m = n_s/n_m = 1{\cdot}452/1{\cdot}333 = 1{\cdot}09$. At high m values, small ripples appear on the sides of the K curves but are not observed for m values as low as 1·1. In the most exact calculations, variations in n_m and n_s with both wavelength and temperature must be allowed for.

3. Determination of mean globule diameters in homogenized emulsions

Using the calculated K values, Eqn (7) can be used to demonstrate the effects of changes in globule diameter on the $E^{1\%}_{1\,cm}$ values at a series of wavelengths. Fig. 1 shows the curves calculated for wavelengths of 0·5, 1·0 and 1·5 μ, for an m value of 1·09.

As can be seen from the curves in Fig. 1, the turbidity is relatively insensitive to the small globules and, since the large globules are reduced in size during the homogenization process, homogenized emulsions approximate to monodisperse systems from a spectroturbidimetric standpoint. This approximation becomes better at higher homogenization pressures where the globule diameters are smaller.

Single-wavelength method

Where the concentration of the suspended phase is already known, curves similar to those of Fig. 1 can be used to read-off d values from the $E^{1\%}_{1\,cm}$ values found by experiment. Measurements made at shorter wavelengths are more sensitive to changes in globule sizes, provided the diameters are less than about 2 microns. For larger globule diameters, $E^{1\%}_{1\,cm}$ is multivalued so that it is preferable to use the longer wavelengths for emulsions containing slightly larger globules. As shown by Bateman, Waneck and Eshler (1959) in their

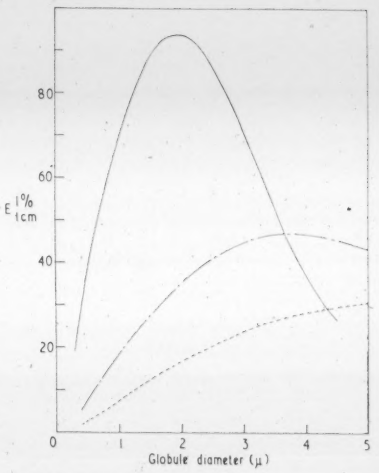


Fig. 1. $E^{1\%}_{1\,cm}$ as a function of globule diameter at wavelengths 1·5 μ (– – –), 1 μ (— — — —), 0·5 μ (—). $m = 1{\cdot}090$.

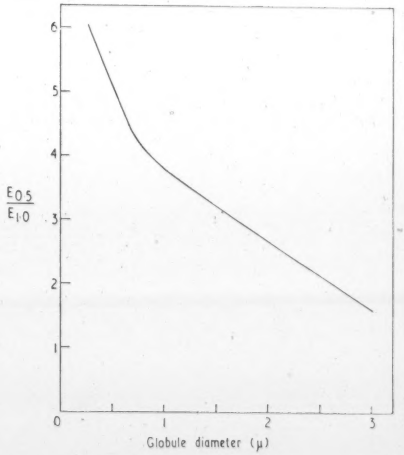


Fig. 2. Ratio of $E^{1\%}_{1\,cm}$ values at wavelengths 0·5 and 1·0 μ as a function of globule diameter. $m = 1{\cdot}090$.

detailed studies of the spectroturbidimetry of polystyrene latex emulsions, the single-wavelength method provides the most accurate d values.

Double-wavelength method

In some cases, the concentration of the suspended phase is unknown so that d must be evaluated by a less accurate method using extinctions measured at two or more wavelengths.

It can be shown that the ratio of $E^{1\%}_{1\,cm}$ values at two different wavelengths is a function of the globule diameter, as illustrated in Fig. 2 for the wavelengths 0·5 and 1·0 μ.

Equations (6) and (7) show that for any particular globule diameter $E_1/E_2 = K_1/K_2 = D_1/D_2$, where the subscripts refer to the wavelengths 1 and 2. Thus for the same emulsion dilution, the ratio of the optical densities at two suitable wavelengths can be used to determine d without a knowledge of the concentration of the suspended phase.

4. Determination of the concentration of the suspended phase

For emulsions in which the particle size distributions are the same, Eqn (5) shows that a linear relationship holds between the optical density and the concentration of the suspended phase, analogous to the Beer's law relationship for absorption. Where possible, it is preferable to use the wavelength of maximum extinction to obtain the greatest concentration sensitivity, coupled with the minimum errors due to variations in wavelength. These conditions apply at the first maximum of the (K, α) curve.

Errors due to small variations in particle size distribution can sometimes be eliminated by suitable choice of wavelength. From the (K, α) curve it is possible to derive a relationship between K/α and α as shown in Fig. 3. Near the maximum

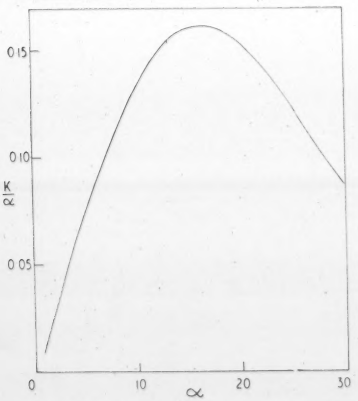


Fig. 3. Calculated relationship between K/α and α for $m = 1{\cdot}090$.

of this curve, K/α can be taken as constant, i.e. $K\lambda/d$ is constant. At a suitable fixed wavelength, Eqn (7) shows that $E^{1\%}_{1\,cm}$ is no longer dependent upon d. Meehan and Beattie (1961) have applied this method to polydisperse silver bromide suspensions where the almost two-fold range in d value was narrow enough for K/α to be regarded as constant. In this case the high m value of 1·75 leading to an α value of about 2 for the $(K/\alpha, \alpha)$ maximum makes the method applicable to suspensions with particle diameters of about $\frac{1}{4}$ micron when visible radiation is used. For oil-in-water emulsions, with an m value of about 1·1, the corresponding globule diameter is about 2 microns for visible radiation.

Most biological and technical emulsions show variations in the size distributions of the suspended globules. When such emulsions are homogenized sufficiently well for spectroturbidimetric studies, the mean globule diameters are usually of the order of one micron and, as can be deduced from Fig. 3, the above technique could only be satisfactorily applied using short wavelength ultra-violet radiation. For the study of such emulsions with visible radiation, the two-wavelength method can be used to obtain a value for d, which then allows the concentration of the suspended phase to be determined from either of the two $E^{1\%}_{1\,cm}$ values. As shown in Fig. 4, it is convenient to prepare a working graph of $E^{1\%}_{1\,cm}$ against the optical density ratio, using a series of diluted homogenized emulsions of known concentrations but of differing and unknown mean globule diameters. This relationship can also be calculated theoretically, enabling a scale of d values to be included. Fig. 4 demonstrates the

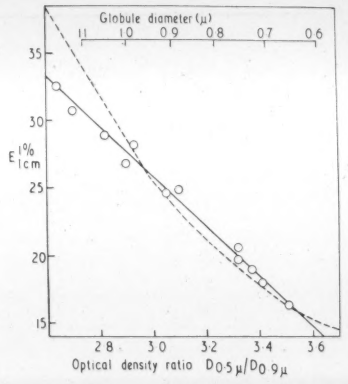


Fig. 4. Relationship between $E^{1\%}_{1\,cm}$ at 0·9 μ and the optical density ratio for ice-cream mixes of $m = 1{\cdot}096$.
○ experimental values, – – – – – calculated relationship.

good agreement obtained between the theoretical and experimental values for a series of homogenized ice-cream mixes in which $m = 1{\cdot}096$.

5. Experimental requirements

The optical requirements for turbidimetric determinations from extinction measurements are more stringent than those for normal absorption measurements. Since commercial spectrophotometers are designed primarily for carrying out absorption measurements with clear solutions, some modifications are normally required before spectroturbidimetric studies can be carried out.

It can be shown from diffraction theory that the radiation scattered by a particle becomes increasingly concentrated in

the forward direction as α and m increase, i.e. as either the particle size or the refractive index ratio increases, or as the wavelength decreases. This is illustrated qualitatively by Fig. 5.

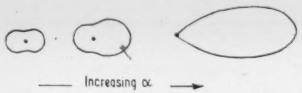


Fig. 5. Effects of $\alpha = \pi d/\lambda_m$ on the distribution of the radiation scattered by a single particle. Incident light beam travels from left to right.

In photometric measurements made with clear solutions, only the radiation which is not attenuated by absorption reaches the detector, but for translucent samples some of the scattered radiation will also reach the detector. Thus the measured optical density will be less than the theoretical value by an amount which depends upon the proportion of scattered radiation reaching the detector. Since the scattered radiation falls off in intensity according to the inverse-square law, a detector placed a long way from the sample will collect a negligible fraction of the scattered radiation and will give a true optical density reading.

Detailed studies of the requirements of optical systems carried out by a number of workers (Lothian and Chappel 1951, Gumprecht and Sliepcevich 1953, Heller and Tabibian 1956, 1957) have shown that the angles subtended at the sample by both the detector and the source should be as small as possible, particularly for high α values. Although specially designed instruments have been constructed by Lewis and Lothian (1954) and by Heller and Tabibian (1957) for studying suspensions of large particles with high refractive indices, i.e. large α values, Bateman, Waneck and Eshler (1959) have shown that a standard Beckman DU spectrophotometer can be modified for studies of the turbidity of polystyrene latex emulsions ($m = 1{\cdot}20$), by interposing a system of lenses and apertures between the sample compartment and the photocell box. For spectroturbidimetric studies on fat emulsions ($m \simeq 1{\cdot}1$), Goulden (1960) has shown that sufficiently accurate optical densities can be obtained at longer wavelengths with a Unicam type SP500 spectrophotometer, provided the photocell box is placed about 10 in. away from the sample compartment.

For most applications of spectroturbidimetry, a narrow spectral bandwidth is unnecessary. A simple filter instrument can therefore be used in place of a more expensive instrument incorporating a monochromator. The effective wavelength of the particular filter-detector combination can easily be determined from an examination of the same emulsion at a definite wavelength with a suitable spectrophotometer. Goulden (1958a, 1960) has described modifications to the Hilger Biochem Absorptiometer which enable turbidity measurements to be carried out.

6. Polydisperse emulsions

In principle, light-scattering techniques can be applied to polydisperse emulsions and it is possible to calculate the turbidity spectrum from a knowledge of the complete globule size distribution. This calculation involves the summation of the turbidity contributions from each globule size group at a series of wavelengths using the values of $\Sigma K_i p_i d_i^2/\Sigma p_i d_i^3$, where p_i is the proportion of globules of diameter d_i having scattering coefficient K_i. Goulden (1958b) has applied this treatment to non-homogenized milk samples and has shown how the spread of globule sizes reduces the customary sharp change in optical density with wavelength to a very flat curve.

From an analytical standpoint, the reverse process of determining a size distribution curve from turbidity measurements is more important and Wallach, Heller and Stevenson (1961) have recently shown that by assuming a standard mathematical form for the distribution curve, various size parameters can be deduced from turbidity measurements made at several wavelengths. Except for the specific m values for which parameters are given, an electronic computer appears to be required for the analysis of the data.

7. Practical applications of spectroturbidimetry

Determination of the fat content of homogenized milk and ice-cream mixes

Pettinati and Haugaard (1959) have devised a rapid spectrophotometric method for the quantitative analysis of homogenized ice-cream mixes, based upon their method for the determination of the fat in homogenized milk (Haugaard and Pettinati, 1959). From measurements made with the detector close to and separated from the diluted emulsion sample, a measure of the mean fat-globule size was obtained and by use of a nomogram, the fat percentage calculated with an accuracy of about $\pm 2\%$. Since the angular changes in the distribution of scattered radiation follow a similar pattern to those produced by change of wavelength, this method is essentially similar to the two wavelength method. Goulden and Sherman (1962) have modified a simple filter absorptiometer for the analysis of ice-cream mixes by means of the two-wavelength method.

Estimation of the mean fat globule diameters in homogenized milk samples

Goulden and Phipps (1960) have shown that the spectroturbidimetric technique provides a simple and rapid method for estimating the mean fat globule diameters in homogenized milk samples. Particularly at higher homogenization pressures, the spectroturbidimetric results are in good agreement with the mean volume-surface diameters $d_{vs} = \Sigma nd^3/\Sigma nd^2$ as obtained by the microscope method. Using the spectroturbidimetric method, it has been shown that d_{vs} is related to the homogenization pressure P by an equation of the form

$$d_{vs} = AP^{-n}$$

where A and n are experimental constants, n having a value of about $0{\cdot}5$.

Globule sizes in emulsions used for intravenous injection

For injection purposes, the globule diameters of oil-in-water emulsions should be about the size of the *chylomicrons*, i.e. less than about $0{\cdot}5\ \mu$. Phipps (1960) has designed a small-capacity laboratory homogenizer for the preparation of such emulsions. Although the globule sizes of the emulsions prepared in this way were too small to be measured by the microscope method, the spectroturbidimetric method was able to show that the fat globule diameters were about $0{\cdot}25\ \mu$.

Analysis of polystyrene latexes

Bateman, Waneck and Eshler (1959) have shown how a standard spectrophotometer can be modified for spectroturbidimetric analysis of plastic emulsions. Several different procedures were described and it was shown that the particle

diameters obtained from $E^{1\%}_{1\,cm}$ values agreed to within 1% of the values determined by electron microscopy. An accuracy of better than 6% was claimed for the spectroturbidimetric method for determining total solid contents.

Estimation of bacterial cell counts

Although not strictly emulsions, suspensions of bacterial cells are frequently examined by turbidimetric and nephelometric methods. In a similar manner to emulsions, the turbidity of a suspension of cells will be dependent upon the cell concentration, the cell sizes and refractive index, apart from possible instrument parameters. A further complication is introduced by the direction of orientation of the cells when they are not true spheres. All these factors tend to reduce the accuracy of spectroturbidimetric studies and, in practice, it is usual to construct experimental calibration curves from samples with known cell counts. During the growth of an organism, the size and possibly the refractive index of the cells may change so that the experimental conditions must be carefully standardized.

Acknowledgments

The author thanks Mrs. J. S. M. Conway for help in experimental work and calculations and Mr. L. W. Phipps and Mr. G. F. Lothian for valuable discussions and advice.

References

BATEMAN, J. B., WANECK, E. J., and ESHLER, D. C., 1959, *J. Colloid Sci.*, **14**, 308.

GOULDEN, J. D. S., 1958a, *Dairy Ind.*, **23**, 558.

—— 1958b, *Trans Faraday Soc.*, **54**, 941.

—— 1960, *J. Dairy Res.*, **27**, 67.

GOULDEN, J. D. S., and PHIPPS, L. W., 1960, *Proceedings of the 3rd International Congress for Surface Activity* (Universitätsdruckerei Mainz GMBH), **C13**, 190.

GOULDEN, J. D. S., and SHERMAN, P., 1962, *J. Dairy Res.*, in the press.

GUMPRECHT, R. O., and SLIEPCEVICH, C. M., 1951, *Tables of Scattering Functions for Spherical Particles* (University of Michigan: Engineering Institute).

—— 1953, *J. Phys. Chem.*, **57**, 90.

HAUGAARD, G., and PETTINATI, J. D., 1959, *J. Dairy Sci.*, **42**, 1255.

HELLER, W., 1957a, *J. Chem. Phys.*, **26**, 920.

—— 1957b, *J. Chem. Phys.*, **26**, 1258.

HELLER, W., and MCCARTHY, H. J., 1958, *J. Chem. Phys.*, **29**, 78.

HELLER, W., and PANGONIS, J., 1957, *J. Chem. Phys.*, **26**, 498.

HELLER, W., and TABIBIAN, R. M., 1956, *J. Colloid Sci.*, **11**, 195.

—— 1957, *J. Colloid Sci.*, **12**, 25.

VAN DE HULST, H. C., 1957, *Light Scattering by Small Particles* (New York: John Wiley).

LEWIS, P. C., and LOTHIAN, G. F., 1954, *Brit. J. Appl. Phys., Suppl.* No. 3, S71.

LOTHIAN, G. F., and CHAPPEL, F. P., 1951, *J. Appl. Chem.*, **1**, 475.

MEEHAN, E. J., and BEATTIE, W. H., 1961, *Analyt. Chem.*, **33**, 632.

MIE, G., 1908, *Ann. Phys. Lpz.*, **25**, 377.

PETTINATI, J. D., and HAUGAARD, G., 1959, *Ice Cr. Fld.*, **74**, 44.

PHIPPS, L. W., 1960, *Lab. Pract.*, **9**, 313.

WALLACH, M. L., HELLER, W., and STEVENSON, A. F., 1961, *J. Chem. Phys.*, **34**, 1796.

Photovoltaic response of selenium barrier layer cells to X-radiation in the energy range of 15–40 keV

by R. FEINBERG,* Dr.Ing., M.Sc., M.I.E.E., and G. E. RHEAD,† M.Sc.Tech., Grad.Inst.P., Electrical Engineering Department, The Manchester College of Science and Technology, Manchester 1

MS. received 9th March 1961, in revised form 24th April 1961

Abstract

Twenty selenium barrier layer cells, made by a modern commercial process with a variety of modifications, were examined with irradiation of x-rays in the 15–40 kev energy range and the behaviour was compared with their response to light. The x-ray sensitivity is found to be independent of intensity of irradiation at constant x-ray tube anode voltage, but linearly proportional to anode voltage at constant anode current. The form of transient response to x-irradiation depends on the intensity of irradiation and the previous history of cell irradiation, a time of rest of about 2 hours being required to achieve the 'first-irradiation' response. The fatigue effect does not occur in the steady-state values to photovoltage and photocurrent. There is some evidence for a qualitative relationship between x-ray and optical photovoltage sensitivities of a cell, but there is no relationship between the photocurrent sensitivities.

1. Introduction

A PHOTOELECTRIC response of cuprous oxide and selenium barrier layer cells to 40 kev x-rays was briefly reported by Lange and Selényi (1931); Scharf and Weinbaum (1933) observed a fatigue effect in a selenium cell exposed to x-rays of 40–140 kev; Gleason (1933) found, in work on a number of selenium cells with x-rays up to 600 kev, that cells having equal sensitivities to light do not necessarily have equal sensitivities to x-radiation, and further that the time lag of response to x-rays was of the order of 15 s, whereas that to light giving the same photocurrent was about 6 s. Evidence of response to soft x-radiation was produced by Felsinger (1937) by measurements with x-rays of 5–25 kev on single cuprous oxide and selenium cells, and by Sandström (1938) with x-rays of 1·5, 2·5 and 3·5 kev on two selenium cells of different makes. Blet (1953) made some measurements on a single selenium cell with x-rays of 20, 40, 60 and 80 kev in connection with theoretical deductions on quantum efficiency.

In a search for a simple method to measure the intensity of x-radiation in the lower kev range, it was decided to examine systematically the photovoltaic performance of a variety of selenium cells produced commercially by a process which is based on an improved technique (Preston 1950). The energy range of x-radiation was 15–40 kev.

* Seconded from Ferranti Ltd.

† Now at the Department of Natural Philosophy, University of Glasgow.

2. Experimental details

Figure 1 illustrates the principle of design of the selenium cells used in the investigation. The metal ring 4 constitutes

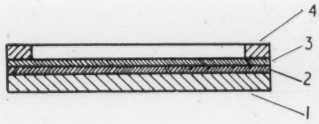


Fig. 1. Principle of construction of a selenium barrier layer cell.

(1), circular steel base plate, 45 mm diameter, 1 mm thick; (2) thin even coating obtained from pure molten selenium subjected to pressure and heat treatment; (3) optically transparent metal layer containing a thin gold film on some metal oxide; (4) ring electrode of a metal with low melting point.

the negative electrode and the steel base 1 is the positive electrode. The free surface of the layer 3 is lacquered to protect it against any atmospheric effects. The photosensitive surface of the cell has a diameter of 40 mm.

Altogether, twenty cells, numbers 1 to 20 in Table 1, were examined, the cells numbered 1–5 having been selected as optical sub-standards from various production batches, and the remainder having been specially prepared in two different batches, numbers 6–11 and 12–20, respectively, with several modifications as listed in Table 1.

For measurement each cell was screened electrically against any effects from the high-voltage electric field of the nearby x-ray equipment, and screened optically against light by placing it inside a metal block connected to earth, with the front covered with aluminium foil of a surface density of 4·85 mg/cm^2. An 'Electrothermal' (Electrothermal Engineering Ltd.) heating tape was wrapped around the metal block to increase the temperature of the photocell. An S.T.C. thermistor type M53 was placed directly behind the steel base of the cell to measure the temperature. Changes of temperature to the nearest 0·2 deg C could be observed.

The source of x-radiation was a low-power sealed-off x-ray tube, Ferranti type B110, which has a triode structure with a copper target rated at a maximum dissipation of 25 w, and a directly heated cathode. The tube was operated with direct current at anode voltages up to 40 kv and it was possible to keep the anode current stable to about 0·01 ma at its maximum value of 1 ma. The dose rate of x-radiation generated was calibrated with an ionization chamber mounted with the centre of the effective air volume 60 mm from the anode of the x-ray tube. Fig. 2 gives the intensity of radiation

Table 1. *Cells investigated and results obtained*

Cell No.	Production technique	X-ray sensitivity: Photocurrent ($\mu A r^{-1} s^{1}$)	X-ray sensitivity: Photovoltage ($mV r^{-1} s^{1}$)	Optical sensitivity: Photocurrent ($\mu A/lx$)	Optical sensitivity: Photovoltage (mV/lx)
1	Standard	0·61	1·03	0·41	0·56
2	Standard	0·17	0·04	0·15	0·024
3	Standard	0·35	0·51	0·45	0·60
4	Standard	0·37	0·26	0·35	0·22
5	Standard	0·35	2·20	0·46	0·64
6	Surface unlacquered	0·80	1·39	0·30	0·17
7	Thicker selenium layer	0·09	13·6	0·20	0·07
8	Metal oxide deposited for 7 s	0·97	24·6	0·52	2·08
9	Metal oxide deposited for 14 s	0·86	45·0	0·48	2·92
10	Metal oxide deposited for 28 s	0·62	49·0	0·50	2·90
11	Metal oxide deposited for 56 s	1·00	64·3	0·51	2·85
12	Metal oxide deposited for 7 s	0·43	13·6	0·54	2·10
13	Metal oxide deposited for 56 s	0·51	21·1	0·55	2·10
14	Metal oxide deposited for 70 s	0·56	29·6	0·58	2·50
15	Metal oxide deposited for 85 s	1·17	31·3	0·49	1·81
16	Metal oxide deposited for 100 s	0·49	23·6	0·50	3·45
17	No gold film, metal oxide deposited for 56 s	0·79	47·5	0·49	3·40
18	No gold film, metal oxide deposited for 70 s	0·73	41·4	0·59	3·33
19	No gold film, metal oxide deposited for 85 s	0·82	33·0	0·57	3·00
20	No gold film, metal oxide deposited for 100 s	0·71	30·0	0·59	3·33

of the tube in terms of dose rate D, in röntgens per second, at an anode current $I_a = 0{\cdot}50$ mA and as a function of anode voltage V_a; D is virtually a linear function of V_a in the range 15 to 40 kv.

The photocell and x-ray tube were mounted at a distance of 60 mm ± 1 mm between the centres of the photocell and x-ray tube anode. In order to vary the sensitive area of a photocell to x-irradiation, apertures of various sizes, made

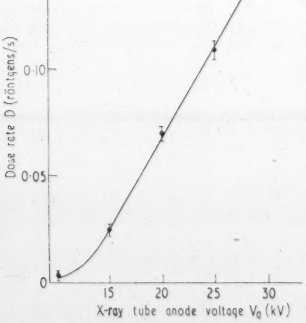


Fig. 2. The dose rate of the x-ray tube.

in lead disks 2 mm thick, were placed in turn directly in front of the photocell mounting. For qualitatively testing the effect of the aluminium foil screen on the x-ray sensitivity of a photocell, filters of various thicknesses were made employing different numbers of layers of the same type of aluminium foil as used for the screen.

3. Experimental results

The cells examined were measured, where not stated otherwise, for photocurrent I_s, i.e. the current theoretically into zero resistance, but practically into a resistance equal to the galvanometer coil resistance value of 154 Ω, and for photovoltage V_0, i.e. the open circuit e.m.f. For a given level of illumination maximum deflection was reached in about four seconds when a cell was suddenly exposed to light, but the response to sudden exposure to x-irradiation showed a complex pattern.

Figure 3 illustrates the response of cell No. 5 to sudden exposure to constant intensity of x-irradiation for various

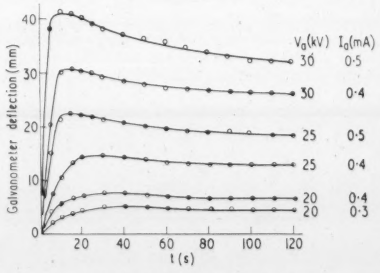


Fig. 3. Response of cell No. 5 to sudden exposure to constant intensity of x-irradiation.

values of x-ray tube anode voltage V_a and anode current I_a, i.e. for various values of intensity of steady x-irradiation. The photocurrent i rises gradually, following a curve of the type

$$i = A\{1 - \exp(-t/T_A)\} - B\{1 - \exp(-t/T_B)\}$$

with $A > B$ and $T_A < T_B$. The value of $A - B$ is a measure of the intensity of x-irradiation; T_A appears to decrease

when the intensity of x-irradiation increases. At a given constant intensity of x-irradiation, steady-state values of photocurrent and photovoltage response are attained in about 2 minutes for lower levels of x-irradiation and in more for high levels.

Figure 4 demonstrates for one of the x-ray intensity values the effect of previous irradiation on the shape of the response

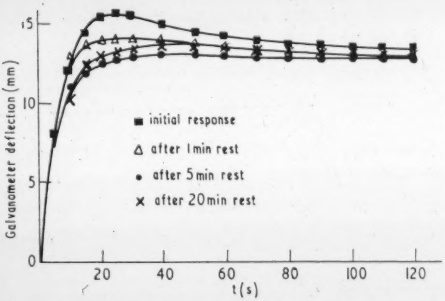


Fig. 4. Effect of previous x-irradiation on response of cell No. 5.

curve after different intervals of rest, each irradiation period lasting altogether 2 minutes. As the rest period is increased the photocurrent response first gradually approaches a curve of the type

$$i = C\{1 - \exp(-t/T_C)\}$$

where $C = A - B$, as above, and T_C is probably equal to T_B, the value of T_A in the relation above having gradually approached that of T_B as the intervals of rest have increased. When the intervals of rest exceed a certain critical value of time the curves of photocurrent response gradually reverse to their original shape, which they regain after an interval of rest of about 2 hours.

When x-irradiation was interrupted the photocurrent decayed rapidly at first, in 5 seconds to less than 5% of the magnitude before interruption and in 15 seconds to less than 0·1%.

The photovoltage appeared to exhibit similar response characteristics. Since no electrometer valve was used at this stage, measurement of the photovoltage required a manual balancing of a bridge arrangement, and it was not possible to measure a photovoltage quickly enough to plot accurately any curves giving reliable information.

The steady-state values of photocurrent I_s and photovoltage V_0 of cell No. 5 as functions of x-ray tube anode current I_a and anode voltage V_a were found to be linear functions of I_a, namely, $I_s = k_i I_a$ and $V_0 = k_v I_a$, the coefficients k_i and k_v being dependent on V_a. For corroboration, a similar set of characteristics was obtained showing I_s and V_0 as functions of V_a, with I_a as parameter, from which the approximate empirical relationships $I_s = m_i(V_a - 11)^2$ and $V_0 = m_v(V_a - 11)^2$ were derived, with m_i and m_v being dependent on the magnitude of I_a. From values of I_s and V_0, with $I_a = 0{\cdot}50$ mA in both sets of characteristics, and the calibration curve of Fig. 2, the x-ray sensitivity curve (Fig. 5) for photocurrent, $s_i = I_s/D$, and for photovoltage, $s_v = V_0/D$, was derived.

The x-ray sensitivities of all cells were determined at $V_a = 30$ kv and $I_a = 0{\cdot}50$ mA. Their values are tabulated in Table 1. For comparison, the optical sensitivities of the cells, measured at an illumination of 72 lux and expressed as

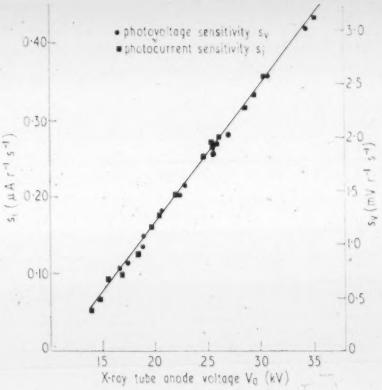


Fig. 5. x-ray sensitivity of cell No. 5 as a function of x-ray tube anode voltage for $I_a = 0{\cdot}50$ mA.

photocurrent and photovoltage per unit intensity of illumination, i.e. in μA/lux and mV/lux, respectively, are also given.

Typical of the effect of temperature on x-ray sensitivity of the various photocells is the characteristic of cell No. 5, presented in Fig. 6. Steady-state photocurrent I_s and photo-

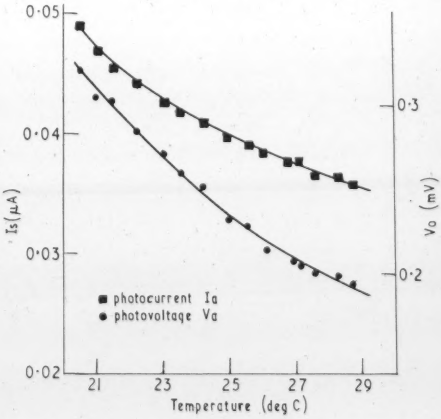


Fig. 6. Effect of temperature on x-ray response of cell No. 5 at $V_a = 30$ kv, $I_a = 0{\cdot}50$ mA.

voltage V_0 were measured in the temperature range 20–30° C. A substantially lower magnitude of temperature effect was found in the optical response of the cells.

Varying the load resistance R_L of a cell at a steady level of x-irradiation gave a photocurrent which changed little when

$R_L < 500\,\Omega$, but decreased substantially as R_L was further increased. Keeping R_L constant at various values ranging from $154\,\Omega$ to $10154\,\Omega$, and altering the intensity of x-irradiation by changing the x-ray tube anode current I_a at a constant value of anode voltage V_a, gave response curves with a linear relation between I_s and I_a starting at the origin.

Changing the effective area of x-irradiation with apertures of various sizes in lead disks showed that the photocurrent changed linearly with the size of irradiated area, provided that the latter remained within the boundary of the region of uniform intensity of the x-radiation beam.

Placing aluminium foil filters of various thicknesses in the x-ray beam in front of a photocell, in addition to its aluminium foil screen, gave an attenuation of photocurrent response, the relative magnitude of attenuation increasing with increasing filter thickness but remaining practically independent of the magnitude of x-ray tube anode voltage V_a when 25 kv$\leqslant V_a \leqslant$40 kv. However, for 25 kv$> V_a >$15 kv the values of relative attenuation gradually increased as V_a was reduced. Table 2 lists the magnitudes of the effect of filters on the photocurrent response for 25 kv $\leqslant V_a \leqslant$ 40 kv, the thickness of the filters being expressed in terms of the number of aluminium foil layers used for a filter.

Table 2. *Effect of aluminium foil filters on photocell response*

Filter thickness (layers)	1	3	7
Attenuation (%)	12	18	25

4. Conclusions

(1) The 'fatigue' effect in the response of a cell to x-irradiation is shown to manifest itself only in the transient state, but not in the steady-state values of photocurrent and photovoltage. The cause of the fatigue effect was not searched for. Perhaps it may be the consequence of some electron trapping mechanism in the selenium or barrier layer of a cell.

The steady-state values of photocurrent and photovoltage at a given intensity of x-irradiation are virtually independent of the previous history of x-irradiation of a cell.

(2) The negative temperature coefficient of x-ray photocurrent and photovoltage sensitivities of a cell is of the order of several per cent per deg C and is substantially higher than that of the optical sensitivity of the cell. Further work is needed to find the reason for the negative temperature coefficient, and to test whether there is any relationship to the fatigue effect in the transient component of photocurrent and photovoltage response discussed above.

(3) The x-ray sensitivity of a cell is independent of the intensity of x-irradiation when the spectral characteristic of irradiation remains constant. When I_a = constant the x-ray sensitivity of the cell is linearly proportional to V_a for $V_a \geqslant$ 15 kv. The value of 15 kv appears to be incidental to the performance of the particular x-ray tube used.

(4) At different values of photocell load resistance the photocurrent response of a cell to x-irradiation is similar to its optical behaviour at weak illumination.

(5) There is an indication of some qualitative relationship between x-ray and optical voltage sensitivities. This may perhaps be based on some underlying quantitative relation masked by some secondary effects.

The tendency of increased thickness of the metal oxide layer to produce higher values of x-ray sensitivities and also higher values of optical photovoltage sensitivity, at virtually unaltered optical photocurrent sensitivity, may possibly be a consequence of a resultant improvement of the semi-conductor structure between the two electrodes of the cell.

Further tests are needed to elaborate any relationship between x-ray and optical photovoltage sensitivities of a cell.

(6) There is no relation between x-ray and optical photocurrent sensitivities of a cell. This may be caused by a fundamental difference in the underlying mechanisms of electron release by photons of x-radiation or of light. It is generally accepted that on illumination electrons are released in the surface zone of the selenium layer near the barrier layer. No understanding exists yet of the underlying mechanism on x-irradiation of a cell. There may perhaps occur a secondary electron release, or even an avalanche effect, and the barrier layer, the body of selenium and even the steel base of the cell have to be considered in this connection. Further work is needed to study the effect of any change of the chemical nature of the base on the x-ray photocurrent sensitivity of a cell.

Acknowledgments

Gratitude is expressed to Messrs. Evans Electroselenium Ltd. for the gift of the photocells which they prepared, and to the Physics Department of the Christie Hospital, Manchester, for the help received in calibrating the x-ray tube with a particular type of ionization chamber, which they had designed and constructed.

References

BLET, G., 1953, *J. Phys. Radium*, **14**, 368.
FELSINGER, H., 1937, *Ann. Phys., Lpz.*, **29**, 81.
GLEASON, P. R., 1933, *Phys. Rev.*, **43**, 775; **44**, 330.
LANGE, B., and SELÉNYI, P., 1931, *Naturwissenschaften*, **19**, 639.
PRESTON, J. S., 1950, *Proc. Roy. Soc.* A, **202**, 449.
SANDSTRÖM, A. E., 1938, *Phil. Mag.*, **26**, 906.
SCHARF, K., and WEINBAUM, O., 1933, *Z. Phys.*, **80**, 465.

Cathode sputtering in inert-gas glow discharges

by B. J. STOCKER, B.Sc., A.R.C.S., Grad.Inst.P., Mullard Research Laboratories, Salfords, Redhill, Surrey

MS. received 5th April 1961

Abstract

Measurements have been made of the rate of deposition of sputtered molybdenum films in the abnormal glow discharge in inert gases over the pressure range 3–21 mm Hg, as a function of the gas pressure p and the current i through the discharge tube. The rate of sputtering was found to be proportional to $(i/p)^{2.5}$ in both neon and the Penning mixture 99% neon–1% argon. In helium sputtering was negligible, but the addition of only a trace of neon caused appreciable sputtering to occur.

1. Introduction

THE phenomenon of cathode sputtering, the emission of atoms from a target bombarded by positive ions, has been known for over a century since the observations of Grove and Pluecker. Early sputtering measurements were made in the glow discharge at pressures around 1 mm Hg. However, under these conditions the incident ions can have any energy between zero and that corresponding to the cathode fall of potential and the sputtered particles travel from the cathode to a collector by a diffusion process, in which more diffuse back than escape. In addition, the three variables, current, cathode fall of potential and pressure cannot be varied independently of one another. The most extensive measurements in this pressure range were carried out by Guentherschulze (1926).

More recently, sputtering measurements have been extended to lower pressures, where the ion energy can be more readily controlled and the diffusion effect eliminated, so as to obtain more fundamental information of the actual process occurring at the target. This has been achieved by applying a magnetic field to the glow discharge (Penning and Moubis 1940), by immersing the target as a negative probe in an arc discharge plasma, or by bombarding it with an ion beam. The most detailed measurements have been made by Wehner (1956, 1957, 1958) using the second method, with a pool-type mercury arc as the source of ionization. Wehner (1955) has also written the most recent review article on sputtering in which he describes other experiments and the theories of sputtering under these conditions.

As far as is known no sputtering measurements at higher pressures than in the early measurements have been reported, other than those of Rockwood (1941) using commercial tubes with oxide-coated cathodes in an inert gas. Owing to the complexity of the process occurring it is not possible to calculate the quantity of material sputtered, as a function of the current, voltage and gas pressure, from the fundamental or empirical data obtained at lower pressures. However, cathode sputtering in the abnormal glow discharge at pressures around 10 mm Hg, is of considerable importance as a means of processing glow-discharge stabilizer and reference tubes with molybdenum cathodes to give stable characteristics, as suggested by Penning and Moubis (1946) and Jurriaanse, Penning and Moubis (1946).

The purpose of this paper is to describe some measurements of the rate of deposition of sputtered molybdenum films in the abnormal glow discharge in inert gases, over the pressure range 3–21 mm Hg.

2. Experimental method

The usual method of determining the quantity of material sputtered is by measuring the loss in weight of the target or the gain in weight of a collector, the target or collector normally being removed from the tube for the weighings. Neither method is convenient if many sputter measurements are required, and the former is likely to give errors, since an initial sputtering is required to remove contamination from the target surface and gas. (A collector can be protected from sputtered material during this process.)

The method used in these experiments was based on that used by Wehner (1956). The deposition of sputtered material on a glass collector was observed continuously by its optical transmission, and by using only a small portion of the collector at one time many measurements could be made in one tube. The light transmission of the deposited film was measured by a photometer calibrated in terms of optical density D, where $D = \log (I_0/I)$ and I_0 and I are the incident and transmitted light intensities.

To determine the relation between the optical density and the mass deposited per unit area, a preliminary experiment was carried out, in which molybdenum was sputtered on to microscope cover glasses in neon, at a pressure of a few mm Hg. For convenience in changing the cover glass, the tube used had a demountable foot on which the electrodes and a tungsten filament were mounted (Fig. 1). The optical density of the deposit sputtered on to a cover glass was determined in the tube using the tungsten filament as light

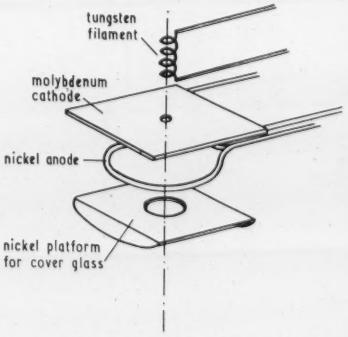


Fig. 1. Diagram showing the positions of the electrodes, filament and cover glass in the tube used for sputtering on to cover glasses.

source, the intensity of which was adjusted to correspond to zero density before sputtering was commenced. After sputtering, the cover glass was removed and the mass of deposited molybdenum determined by chemical analysis. Exposure to air would not affect this method of measuring the mass.

The optical density was found to vary linearly with the mass deposited as can be seen from curve A in Fig. 2.

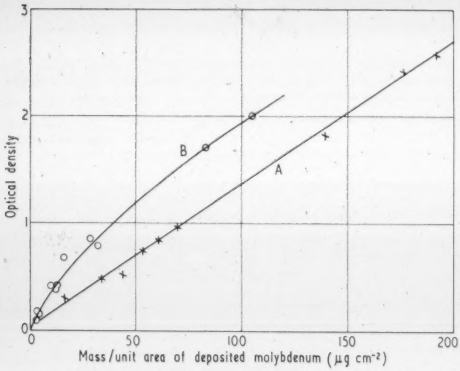


Fig. 2. Variation of optical density with mass of molybdenum deposited per unit area.

Curve B was obtained if the cover slips were removed from the tube for the density measurements, showing the effect of exposure to air. It will be noticed that it is parallel to curve A for thick films.

The measurements of the rate of deposition of sputtered material were carried out in a tube with a molybdenum rod cathode C, surrounded by a concentric cylindrical molybdenum anode A (Fig. 3). A square aperture was cut in the side of the anode through which particles sputtered from the cathode passed to a cylindrical glass collector G, magnetically rotatable so as to enable a series of deposits to be formed. A tungsten filament F mounted inside the anode cylinder was used as a light source. It was sufficiently intense for the light from the discharge to be negligible by comparison. The filament was electrically isolated from the discharge circuit to prevent it from being sputtered.

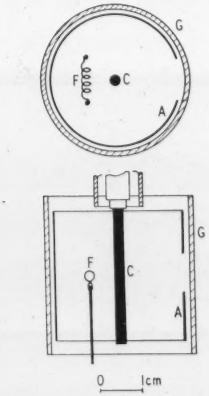


Fig. 3. Diagram showing the positions of the anode A, cathode C, filament F and glass collector G in the tube used for measurements of the rate of deposition.

The tube was sealed to a conventional mercury diffusion pumped system with liquid-oxygen traps, via an Alpert-type bakeable metal tap. A high-pressure (0–25 mm Hg) McLeod gauge was used to measure the pressure of inert gas admitted to the system with an accuracy to 0·1 mm Hg.

The tube was evacuated and baked to 450° C and the metal parts outgassed at a high temperature. During this process the glass collector was slid away from the anode to the other end of the tube. With the glass collector in position the cathode was cleaned by sputtering in inert gas for several hours.

The densitometer was arranged to receive the light from the tungsten filament in the tube via the aperture in the anode, extraneous light being reduced to a minimum. The tube was filled with inert gas at the required pressure, and the discharge current and filament supply were switched on. After allowing the tube to warm up for twenty minutes the gas pressure was measured. The Alpert tap was then closed, thus sealing the tube from sources of contamination in the pump system. The glass collector was rotated to bring an unsputtered portion before the aperture in the anode and the filament current adjusted to give zero density reading. A series of readings of the optical density of the sputtered deposit and the voltage across the tube at constant current were taken at suitable time intervals. A graph of density against time was plotted.

3. Results of the sputtering rate measurements

3.1. *Neon and* 99% *neon–*1% *argon*

A series of graphs of optical density against time was obtained at various currents and pressures in neon and in the Penning mixture 99% neon–1% argon. Typical curves are shown in Fig. 4. These curves were normally straight, except at low densities (probably due to irregularities in the thin films). The slope of the straight portion of each curve was proportional to the rate of deposition, since curve A in

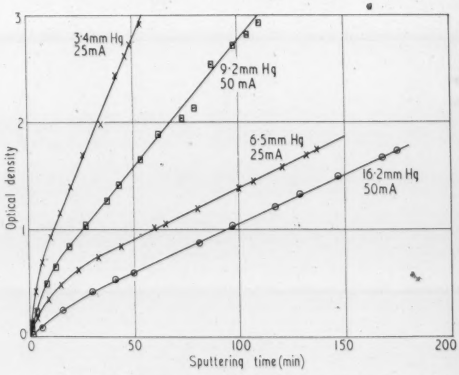


Fig. 4. Typical curves showing the variation of optical density with time in 99% neon–1% argon.

Fig. 2 was linear, and is referred to as the relative sputtering rate. It was not converted to the mass of film deposited on unit area in unit time, since this is still only a relative measure of the rate of sputtering at the cathode (the mass sputtered per unit time). The latter is assumed to be proportional to the rate of deposition, the constant of proportionality being determined by the geometry and angular distribution of the sputtered atoms.

It was found that the sputtering rate could be related to the current and pressure. Fig. 5 shows a series of curves of

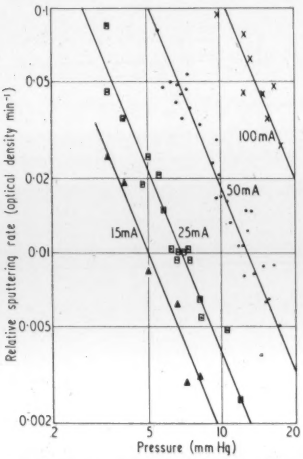


Fig. 5. Variation of sputtering rate with pressure for 99% neon–1% argon.

relative sputtering rate against pressure on logarithmic scales for different currents, the gas being 99% neon–1% argon. It will be noticed that a series of parallel straight lines is obtained of negative slope of magnitude about 2·5. Fig. 6, for neon, is similar, the slopes being approximately the same as in the gas mixture. In Fig. 7 the rate of sputtering at 10 mm Hg is plotted against current on logarithmic scales, both the 99% neon–1% argon and the neon points lying on one line of positive slope with the same magnitude as before. Hence the rate of sputtering M can be related to the current i and pressure p by an empirical expression of the form

$$M = C\left(\frac{i}{p}\right)^{2\cdot5} \qquad (1)$$

where C is a constant which is approximately the same for both neon and 99% neon–1% argon.

The fact that the rates of sputtering in both neon and the Penning mixture are approximately the same is surprising, since in the latter case the majority of the ions are of argon. However, the lower voltage required in the gas mixture, at a given current and pressure, may compensate for the greater mass of the ions.

3.2. *Helium*

An attempt was made to observe sputtering in helium. This proved to be extremely slow and probably due to impurities, since it did not occur at a constant rate, no sputtering at all being detected in the first few hours. However, the addition of only 0·085% of neon increased the rate of sputtering to a value comparable with that in pure neon. This is confirmed by the experiments of Guentherschulze

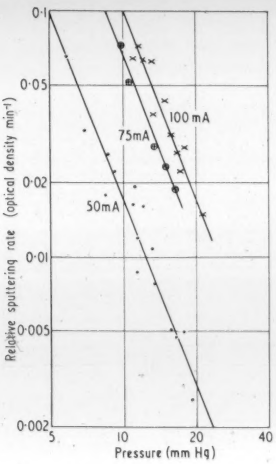


Fig. 6. Variation of sputtering rate with pressure for neon.

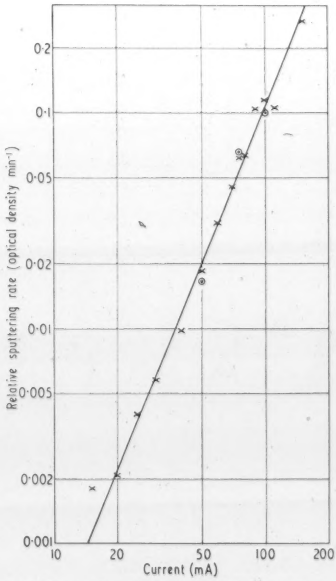


Fig. 7. Variation of sputtering rate with current, at a pressure of 10 mm Hg.
× 99% neon–1% argon. ⊙ neon.

and Meyer (1930) who could detect no sputtering in a helium arc discharge at energies up to 800 v.

4. Discussion

Sputtering measurements in the glow discharge at pressures around 1 mm Hg have frequently shown the quantity of material Q sputtered, per unit charge passed through the tube, to be related to the cathode fall V by a relation of the form

$$Q = \frac{M}{i} = A(V - V_0) \qquad (2)$$

where A and V_0 are constants. Guentherschulze (1926) using silver sputtered in hydrogen, with parallel-plate geometry, showed his results could be fitted either to Eqn (2) or the following equation involving the pressure p and electrode separation d

$$Q = \frac{M}{i} = B\frac{V}{pd} \qquad (3)$$

where B is a constant.

An attempt was made to express the results of the experiments described in this paper by an equation similar to (2) or (3). Although Q did increase with V the fit to either of these equations was poor.

It is not surprising that Eqns (2) and (3) could not be applied satisfactorily. Whereas in the experiments of Guentherschulze the pressure range was 0·3 to 2·2 mm Hg and the voltage range 550 to 1740 v, in these experiments the pressure range was 3 to 18 mm Hg and the voltage range 146 to 340 v in 99% neon–1% argon, and 6 to 21 mm Hg and 164 to 331 v in neon. As a result a much greater proportion of the ions will arrive at the cathode with energies below, or only just above, the threshold energy for sputtering. Also there will be greatly increased back diffusion of the sputtered material.

References

GUENTHERSCHULZE, A., 1926, *Z. Phys.*, **38,** 575.

GUENTHERSCHULZE, A., and MEYER, K., 1930, *Z. Phys.*, **62,** 607

JURRIAANSE, T., PENNING, F. M., and MOUBIS, J. H. A., 1946, *Philips Res. Rep.*, **1,** 225.

PENNING, F. M., and MOUBIS, J. H. A., 1940, *Proc. K. Ned. Akad. Wetensch*, **43,** 41.

—— 1946, *Philips Res. Rep.*, **1,** 119.

ROCKWOOD, G. H., 1941, *Trans Amer. Inst. Elect. Engrs*, **60,** 901.

WEHNER, G. K., 1955, *Advances in Electronics and Electron Physics*, Vol. 7, p. 239 (New York: Academic Press).

—— 1956, *Phys. Rev.*, **102,** 690.

—— 1957, *Phys. Rev.*, **108,** 35.

—— 1958, *Phys. Rev.*, **112,** 1120.

Contents

PAGE

Analysis

Physical methods in chemical analysis 470

Electrical, electrostatic and electronic devices and properties

Analogue computers 470
Lichstrahl-Oszillographen 471
Werkstoffe für elektrische Kontakte 471
Elsevier monographs: Electrostatic separation of mixed granular solids 471
Lectures on communication system theory 472

General

A short history of technology 472

Heat, thermoelectricity, infra-red, radiation, etc.

Kinetics, equilibria and performance of high temperature systems—Proceedings of the First Conference Western States Section, the Combustion Institute 1959 473
Elektrothermie 473
Thermoelectricity 473
Infrared methods: principles and applications 473

Interferometry

Interferometry—National Physical Laboratory Symposium No. 11 474

Mechanics, properties of matter, vibrating systems, elasticity

Mechanics and properties of matter 474
The mechanics of vibration 474
An introduction to the theory of vibrating systems. . . . 475
An introduction to applied anisotropic elasticity 475
Large elastic deformations and non-linear continuum mechanics 475
Impact: the theory and physical behaviour of colliding solids 476
Mechanical waveguides 476

PAGE

Non-destructive testing, ultrasonic devices, X-ray microscopy, etc.

Techniques of non-destructive testing 476
Ultrasonics 477
X-ray microscopy. Cambridge monographs on physics . 477
Progress in semiconductors 478

Nuclear physics

Progress in nuclear energy, series IV. Technology engineering and safety. Vol. III 478
Kernenergie Technik 479
Fast neutron physics: Part I. Techniques 479
Controlled thermonuclear reactions—an introduction to theory and experiment 480
Materials for nuclear engineers 480
Nuclear propulsion 480
Eléments de physique nucléaire 481
Radioisotope laboratory techniques 481

Radioastronomy

The radio noise spectrum 481

Theoretical physics

Small particle statistics 481
Field computations in engineering and physics . . . 482
Plasma physics 483
Wave propagation in a turbulent medium 483
Einführung in die Theorie der Abtastsysteme 483

Vacuum physics

Vaccum technology transactions. Proceedings of the Sixth National Symposium 483
Molecular distillation 484

New books section

Analysis

Physical methods in chemical analysis. Vol. 1. 2nd revised edn. Edited by WALTER G. BERL. (London and New York: Academic Press, 1960.) Pp. xiv + 686. Price £6 15*s*. 9*d*.

The first edition of this book appeared in 1950 and reviews were published in various journals, including *Science Progress*, 1951, **39** (153), 173; *The Analyst*, 1950, **75**, 691; *Analytica Chimica Acta*, 1951, **5** (3), 334; *Analytical Chemistry*, 1950, **22** (7), 956.

This is the first of three volumes. The original intention was to include in it those techniques dependent upon electromagnetic radiation and in the second volume those dependent on electrical measurements. It was subsequently found necessary to publish a third volume dealing with later techniques. The topics dealt with in Vol. I are as follows: Absorption phenomena of x-rays and γ-rays, X-ray diffraction methods as applied to powders and metals, X-ray diffraction as applied to fibres, spectrophotometry and absorptiometry, Emission spectrography, Infra-red spectroscopy, Raman spectra, Refractive index measurement, Mass spectrometry, Electron microscopy and Electron diffraction. It is inevitable that certain recently developed but closely related topics have been relegated to Vol. III, such as X-ray fluorescence, Flame photometry, Microwave spectroscopy and Nuclear magnetic resonance.

The rapid rate of development of even well-established methods has now necessitated the publication of a second edition of the first volume, and it is interesting to note the degree to which the various authors have revised their contributions. For example 'Emission spectrography', which has a strong metallurgical bias, remains sensibly unaltered, and one is left with the feeling that too narrow a viewpoint has been adopted for a book of this character; of some 28 references only one is post 1950. Other topics which have undergone little revision are Infra-red spectroscopy (and this has also suffered from the omission of the excellent diagrams of various spectrometers, which appeared in the first edition), Raman spectra and Refractive index measurement. The last-mentioned is the most poorly illustrated section of the whole book, with only two figures in 45 pages. On the other hand the X-ray diffraction sections, Electron microscopy and Spectrophotometry are all lavishly illustrated. The most drastically revised section is Mass spectrometry, and this has been completely rewritten by C. F. Robinson; of the 393 references quoted over 300 are post 1950. Other topics have also been extensively revised and some have been expanded, so that one is left with the overall impression that, just as the first edition was welcomed and widely appreciated, the publication of the second edition is amply justified and will surely be just as heartily welcomed.

It would doubtless be possible for the critical expert to find fault with individual chapters. But in fact no claim is made that each is comprehensive, nor is that the intention, since specialized textbooks are available. It is, however, justifiably claimed that each subject is dealt with in a manner which will be satisfactory and informative to the general analyst, and the amount of information contained in this volume in fact represents extremely good value for the buyer.

The general style, method of presentation and quality of printing and binding are of the high standard associated with the publishers and, although the price suggests that it is not the sort of book that the young analyst or student will be able to afford, nevertheless it is a 'must' for every library and for every serious worker in the field of analysis or of physics as applied to analytical chemistry. F. W. J. GARTON

Electrical, electrostatic and electronic devices and properties

Analogue computers. By I. I. ETERMAN. (London, New York, Paris: Pergamon Press, 1960.) Pp. ix + 264. Price 50*s*.

This work is an edited English translation of a book first published in Moscow in 1957. The preface explains that the 'principles of action' of analogue computers have been examined both in Soviet and foreign literature whereas the theoretical problems, the preparation of such problems for solution, methods of solution and the methods for the estimation and control of accuracy, have been dealt with less adequately. The aim of the book is thus to deal with these somewhat neglected topics, and they are discussed in the following manner.

The first chapter points to the sort of dynamic system which lends itself to analogue treatment and introduces the Cauchy problem and the Vallé-Poussin theorem. Chapter II reviews the various types of computing units—amplifiers, function generators and the like, which form the basis of analogue computing instruments and machines. The next chapter describes some Russian computers and deals with 'operational procedures' which are, in fact, a series of useful mathematical tricks or devices which are employed to facilitate certain types of problem solution by analogue means. Chapter IV gives the mathematical background to the solution of systems of linear and differential equations and introduces finite difference techniques in these connections. The last chapter is a mathematical appraisal of the errors arising in the solution of linear and non-linear systems of equations, and is intended as a guide to the interpretation of analogue solutions.

There is a sizeable appendix which contains finite difference information, tables of coefficients, and various solutions in tabular form.

This book is curiously unbalanced mainly because it tends to be both a textbook and a reference book at the same time. The result is that in some parts quite important aspects of analogue computer practice are skipped over or even ignored completely, while in others, particularly the last chapter, less common points are treated exhaustively. For example no mention is made of servo-mechanical multiplication or resolvers, and while the error analysis of computer solutions is treated very fully, no mention is made of the inevitable and sometimes very disturbing errors inherent in electronic summation or integration. The book treats the analogue machine as a purely mathematical tool and does not consider simulation as such. Unavoidably perhaps, a good deal has been lost in translation and one frequently finds that the precise scientific sense of a sentence is missing. However, the meaning is usually apparent after a little rephrasing.

Nevertheless, Eterman's work, or rather the part of it which has been translated for this book, contains much interesting material particularly for people using or wanting to use medium-sized analogue computers and similar devices; it contains some material which has not been published before, and the mathematics well repays careful reading. The brief descriptions of Russian general purpose equipment

are interesting even though they are pre-1957 vintage, and one wonders if the extraordinary progress in electronic computing machinery so apparent in Britain and the U.S.A. in the past four years has been matched in Russia. The book's price is not high for what it contains.

G. BLACK

Lichtstrahl-Oszillographen. By W. HÄRTEL, J. DEGENHART, A. KÜBLER, C. SÖRENSEN and J. TRÖGER. (Munich: Oldenbourg, 1961.) Pp. xii + 462. Price DM 64.

After a short introduction to the principles of oscillographs and oscillography the authors go more deeply into theory and practice of vibrator design and deal with sensitivity, frequency response and damping. This includes electrodynamic damping as well as that by means of modern silicone fluids. Special vibrators for power measurements as well as the use of Hall effect generators for this purpose are described.

Circuits and formulae for range extension occupy 13 pages. A 26-page chapter on optical problems follows. This includes the design of mirror drums for visual observation. Various aspects of photographic recording and paper transport are dealt with, especially methods of achieving rapid acceleration of the paper for triggered exposures. Methods of triggering and time markers are described.

A chapter on the historical development of oscillographs especially by Blondel (Paris) and Duddell (The Cambridge Scientific Instrument Co.) leads to a detailed description of instruments by Siemens. This section contains many very clear photographs, drawings and circuit diagrams as well as mathematical treatment of choice of the correct characteristics, possible errors, etc.

There are 64 pages of applications including testing internal combustion engines, torque tests, control circuits, switchgear, etc. A collection of tables giving vibrator characteristics and relationships between errors, damping and frequency ratio is followed by a lengthy bibliography and index.

P. LEWIS

Werkstoffe für elektrische Kontakte. By A. KEIL. (Berlin, Gottingen, Heidelberg: Springer, 1960.) Pp. xi + 347. Price DM 48.

So many new books are devoted to modern technical and scientific developments that it is refreshing to welcome one which deals with a technology which has been with us, in practice, since the beginnings of electricity. The switching of electrical circuits by mechanical means is commonplace, but the potentialities of the technique have received renewed attention in the last ten years, due to increased emphasis on automatic control systems. The few other textbooks on electrical contacts deal, almost exclusively, with the physics of the subject so the present writer's approach from the viewpoint of contact materials is a pleasant change.

Written for the electrical industry the book embodies a critical survey of the present state of knowledge about contact materials. Applications from heavy duty relays to very sensitive micro-switches are covered. The contents are divided into four sections: physical principles, basic metallurgy, material behaviour, and material selection and performance. The treatment of physical principles covers contact constriction theory and arcing phenomena. The reader is then abruptly confronted with over eighty pages listing and describing the general properties of contact materials. The section on material behaviour deals primarily with specific effects associated with the high temperatures invariably encountered with arcing contacts, while the final chapter covers the design of contacts and considers the related topics of springs and solders.

Unless masked by erosion, mechanical wear is encountered in all contact applications. A mere half-dozen papers are devoted to wear processes and a fuller treatment of the subject would have added to the value of the book. Electrodeposited coatings are also dealt with rather superficially in view of their increasing use. In contrast the sections in various sintered materials are extremely well presented. Extensive tables of material properties are an outstanding feature.

The engineer will find the work of value as a survey but whether it will assist him with specific contact problems is doubtful. As the writer points out there is a bewildering variety of possible and actual contact materials and the various physical processes with their attendent chemical and metallurgical effects make the prediction of material performance extremely difficult. In this field where precious metals are so extensively used, economics rather than technical merit is often the major factor in material selection. That the treatment of material selection is not a satisfactory synthesis of the chapters on physical principles and contact materials is no fault of the author who has done an excellent job in presenting this difficult subject.

More attention could have been paid to the compilation of the index but at £4 2*s*. 0*d*. the book, in the light of present standards, is good value.

H. C. ANGUS

Elsevier monographs: Electrostatic separation of mixed granular solids. By O. C. RALSTON. (London: Van Nostrand; Amsterdam: Elsevier, 1961.) Pp. viii + 261. Price 24*s*.

The author of this work is an engineer whose lifetime of experience in the subject is revealed in his descriptions of the different types of equipment, their advantages and limitations. He has also shown his enthusiasm by the amount of work he has undertaken in compiling this book, which, with its 250 references, 350 notes on patents and large number of figures must be a fairly complete study of the field. He fully appreciates the lack of concrete data on performance, which he supplements wherever possible from his extensive personal knowledge of the success or otherwise of the equipment. However, the limited theoretical treatment of the subject is not up to the same standard. In several places the author states his own lack of understanding of the physical principles involved, applauds the valuable part played by the one or two physicists who have studied the subject and stresses the need for more basic work. It is a pity therefore that he introduces the book by attempting a theoretical explanation of the principles of electrostatic separation simplified for the non-physicist. He admits that this approach is not really possible and as a result the first chapter is naïve, confused and technically inadequate. Anyone reading only the first ten or fifteen pages of the book would be liable to condemn it out of hand, which would be most unfortunate in the light of the great practical value of the bulk of the text. The author would have been well advised to have engaged a physicist to write the introductory material and thereby have obtained a more concise and clearer exposition of the principles involved.

Because the physics of the phenomena used are not always understood and because several different mechanisms may be involved in each form of separating equipment, there is considerable difficulty in classification of the methods. Clear distinctions cannot be drawn between types of equipment and the author has been driven to attempt a preliminary breakdown, supplemented by a description of each piece of equip-

ment with an indication of the possible mechanisms that may be operating. There is therefore some overlapping between Chapter III dealing with equipment, and Chapter IV in which applications are discussed. This seems inevitable, since particular forms of equipment are associated mainly with certain types of problem, often merely as a result of historical accident and only occasionally because the problem required a particular type of separator.

The electrostatic method has obviously been used mainly for the separation of minerals and the cleaning of grain or textiles, but occasional uses of it are described covering a wide range of problems including separation according to size, density and moisture content. A striking feature of the uses recorded in the book has been the marked success of failure of apparently similar types of equipment on apparently not very dissimilar materials. The reasons for this inconsistent behaviour have not usually been fully established but are indicated by the author.

Most of the work dealt with appears to have been done in America and Germany, and apart from the allied technique of electrostatic precipitation the principle has not been widely applied in this country, though the causes of this may have been fortuitous; for example the pages on coal cleaning deal with aspects specially relevant to conditions in Germany and America. On the subject as a whole these two countries are credited with 300 patents against a total of 46 for Britain, France and Canada.

The book draws attention to a subject that may not be well known. It supplies a useful source of information and should act as a valuable stimulant to thought for anyone who has a separation problem that is proving to be either difficult or costly. The book is well printed and amply illustrated by clear figures. It may have been just ill-luck that the first few pages of the review copy dropped out as soon as the book was opened.

R. JACKSON

Lectures on communication system theory. Edited by E. J. BAGHDADY. (London: McGraw-Hill, 1961.) Pp. xii + 617. Price 97*s*.

This important work ought to be called more properly the MIT book on communication systems. Its 18 contributors are almost all from the MIT or the Lincoln Laboratory, and the Editor, Dr. Baghdady, is the author of only two of the 23 articles.

First it must be noted with relief that the reader of this book is not overwhelmed by a great mass of unfamiliar mathematics (Lebesgue measure theory, topology, symbolic logic), as has now become customary in MIT publications. The mathematics are either familiar, or are adequately explained. The introduction to probability theory by D. G. Brennan is particularly praiseworthy in this respect. Of the other contributions I should like to single out the two excellent chapters on the representation and the design of signals by R. M. Lerner, perhaps for the selfish reason that it takes as its starting point my work on signal theory of 15 years ago, and makes significant improvements on it.

There are four very thorough chapters on the characterization of noisy channels and methods of improving transmission by W. E. Morrow, H. Sherman, T. Kailath and E. J. Baghdady, which are perhaps the best so far published on this subject. The two chapters on Statistical decision theory, by W. M. Siebert and W. L. Root are as good as the difficult and as yet incomplete subject-matter will allow. J. M. Wozencraft's chapter on digitalized communication in noisy and time-variant dispersive media will be of great interest to practical system-engineers, though somewhat difficult to understand. The other chapters are less systematic, but concern highly important special subjects (Coding by P. Elias, Feedback communication systems by R. P. Rafuse, Microwave applications of semiconductors by A. Uhlir, Masers by R. H. Klingston, Speech systems by I. Pollack, Modulation systems by Baghdady, Long term variability by D. G. Brennan and Communication link design by W. E. Morrow). The last chapter on Satellite Communication by J. B. Wiesner will evoke special interest, by its highly important and topical subject-matter, and also because the opinions of the author, until recently Director of the Research Laboratory of Electronics, and now Scientific Adviser to the President of the United States, carry quite special weight in what promises to become the most exciting adventure in communications since the laying of the first Transatlantic cable.

This is a book which must be recommended to all post-graduate students in communications, to all systems engineers and also to all teachers of mathematics to engineers. It is well worth the moderate price.

D. GABOR

General

A short history of technology. By T. K. DERBY and TREVOR I. WILLIAMS. (Oxford: Clarendon Press, 1960.) Pp. xviii + 782. Price 38*s*.

All those who would like to possess but cannot afford the five-volume *History of Technology* will be grateful to the I.C.I. for having endowed this *Short History*, at a price which everybody can afford.

This is the story of the greatest adventure of mankind, told in unemotional language. I wish I were sure that the reader will be able to supply the emotions, the pride in our ancestors, in those relatively very few who dared to dream of new things which had never existed before, in climates of opinion which punished rather than encouraged the innovator. It is an astonishing story of empiricism, of the thousands of years of fumbling before the advent of organized science, when every fact stood by itself, unrelated to all others. How many unsuccessful experiments must have preceded, for instance, the making of red lead by the Mesopotamians 'by heating lead with basic lead carbonate, itself probably made by addition of natron to a solution of lead salt'? (p. 264). Experimental science could never have started without these old artisans, who supplied the first experimenters with highly developed arts of metal working, glass blowing and lens grinding. The first half of the book, up to 1750, is mostly the story of artisans. The rest brings us up to 1900. 1750–1900 is roughly the first epoch of the co-operation of science and industry, 1900 marks approximately the date from which the artisan became the executive of the scientist, the end of the romantic epoch.

In writing a book such as this the authors were faced by several dilemmas. One was that either they had to explain every technical term for the benefit of the general reader, or to use freely such terms as 'heddle', 'shed-rod' and 'counter-shed', intelligible only to the expert. They have opted generally for the second choice. Others than myself will probably be intrigued when they read that Jeremy Bentham set out to Paris in a 'titiwhiskey' and arrived in a '*pot de chambre*' (p. 212).

Another dilemma is whether a history of technology ought to be illustrated by technological or by historical pictures? The authors prefer the second, which is sometimes regrettable,

because many of the early pictures abound in irrelevancies, such as the garments of the artisan, but leave the reader in the dark about what he is doing. Much would be gained by adding a few schematic diagrams.

I do not wish to carp on the few deficiencies of a book which was so badly needed, and which, on the whole, fulfils its task so admirably. I wish thousands of boys would receive it as prizes in schools, and I wish that thousands of engineers and scientists would keep it at their bedside.

D. GABOR

Heat, thermoelectricity, infra-red, radiation, etc.

Kinetics, equilibria and performance of high temperature systems—Proceedings of the First Conference Western States Section, The Combustion Institute 1959. Edited by G. S. BAHN and E. E. ZUKOSKI. (London: Butterworths, 1960.) Pp. x + 255. Price 80*s*.

The Conference was primarily aimed at those who are engaged in performance calculations for high temperature systems. Groups of papers were devoted to input and output thermodynamic data, computer programming for output thermodynamic data and for standard engine performance parameters. Programmes for combustor and exhaust nozzle design and results of performance calculations were also included.

Elektrothermie. Edited by M. PIRANI. (Berlin: Springer-Verlag, 1960.) Pp. xii + 451. Price DM. 61.50.

Industry's needs for high temperatures (i.e. exceeding 1000° C) are being increasingly met by electrical heating. There are sound economic and technical reasons for this trend. It is therefore timely that a second edition of *Elektrothermie*, which first appeared in 1930, should be published under the editorship of Professor Pirani.

The field covered is wide, and the treatment predominantly technological. Underlying chemical, physical and engineering principles are dealt with sketchily, if at all, but this is no adverse criticism, in a book which aims at presenting the state of the art to those who are interested in a broad non-specialized way. It is unlikely that specialist readers will find here anything that is new to them in their own speciality; but they will probably find stimulus in other chapters.

Individual sections deal with the productions of steel, aluminium, copper, molybdenum, tungsten, titanium, the semi-conductors silicon and germanium, graphite, corundum, the carbides of silicon and boron, calcium silicide, ferro-silicon, phosphorous, cyanamide, vitreous silica and sintered products.

Industrial electric furnaces for temperatures above 1500° C based on resistance, induction or arc heating are described, including a brief note on electrically heated glass tanks. Measurement and control of electric furnaces are briefly but reasonably comprehensively described. Finally, a few examples are given of laboratory applications such as barium carbide production, graphitization and the production of extremely high temperatures, both in the plasma torch and in attempts at nuclear fusion.

The book concludes with a useful bibliography. With one exception (Swiss) it is the work of German authors, and is mainly (but by no means exclusively) concerned with German practice. It is a useful handbook for those interested in the growing field of high-temperature technology.

T. H. BLAKELEY

Thermoelectricity. Edited by P. H. EGLI. (London: John Wiley, 1960.) Pp. x + 407. Price 80*s*.

This book is based on a Conference on Thermoelectricity held in September 1958. It has, however, been expanded and increased in scope in an attempt to meet the requirements of 1960.

The first section introduces the fundamental concepts of thermoelectricity, and includes a particularly stimulating article by Zener. One of the papers in this section refers to thermoelectric refrigeration, but otherwise the main emphasis of the book is on thermoelectric generation. The second section is headed 'Basic Parameters in Thermoelectricity,' but does not in fact deal with these systematically. The first paper gives some general account of thermal conductivity, and the remaining three papers discuss some of the properties of several important thermoelectric materials, including the mixed valency semiconductors.

Section 3 deals with the chemical and physical properties of materials at high temperatures, starting with a paper on the types of compound likely to be stable under these conditions. There is a short paper by Aigrain on the design of materials for thermoelectric generation, followed by a discussion of porous semiconductors and a paper on thermionic-diode generators. The last of these seems distinctly out of place in this book.

The final section deals with the measurement of the properties of thermoelectric materials, particularly at high temperatures. This is probably the most valuable section, as there are formidable difficulties which have not been very fully discussed in the literature.

The book is well produced and most of the articles are well written and presented. Inevitably, there is some overlap and there are some omissions. One criticism is that no attempt has been made to justify what is included or to indicate what is omitted and how important it might be. Thermoelectricity is a subject in which there is a great deal of activity at the present time, some of it empirical and unco-ordinated, and some based insecurely on an inadequate background of the rather complex physics which is involved. The book will help considerably in improving this situation, although more could have been achieved if a paper had been included surveying the different classes of compound so far studied, and indicating the progress made and the problems remaining. The Editor hopes that the book will provide guidance in the development of thermoelectric materials; such guidance as it offers is however rather tentative, and is mainly to be found in a few lines in the two papers by the authors named.

D. A. WRIGHT

Infrared methods: principles and applications. By G. K. T. CONN and D. G. AVERY. (New York: Academic Press, 1960.) Pp. viii + 203. Price $6.80 (54*s*.).

Many scientists now find that infra-red radiation is of value in studying a wide range of problems. Newcomers to this rapidly expanding field are frequently bewildered by the volume of literature on infra-red methods which is spread over a very wide range of journals, all of which are rarely found in a single library. Any manual which deals with the principles and gives some of the applications of infra-red must therefore be warmly welcomed.

The first half of this book describes sources of radiation, optical materials, detectors, amplifiers and dispersive systems. This is followed by sections on applications which include the calibration of detectors, a simple monochromator and instruments for gas analysis and radiation pyrometry. While

the book has been kept to a manageable size, it is inevitable that much of interest and value has had to be omitted from its mere two hundred pages. Thus, for instance, far infra-red techniques and spectroscopic applications are not discussed. Future editions might include a chapter giving some key references to a wider range of applications. On the whole, however, the topics that are dealt with have been well selected. Many research workers who either use infra-red radiation or who want to know more about it will find the book of considerable value. G. R. WILKINSON

Interferometry

Interferometry—National Physical Laboratory Symposium No. 11. (London: H.M.S.O. for D.S.I.R., 1960.) Pp. 471. Price £1 10*s*.

This publication is the record of National Physical Laboratory Symposium No. 11, held at the National Physical Laboratory on 9th–11th June 1959. The great breadth of the field covered is indicated by the session headings, viz. Absolute length measurement and light sources for interferometry; Optical testing and the measurement of relative position; Radio, microwave and intensity interferometry; Spectroscopic and photoelectric interferometry. In addition, there was an introduction session embodying a discourse by Professor Dr. F. Zernike entitled 'Limitations of Interferometry' (not reproduced in this book, as a full treatment is to be given in a later paper) and accounts by Dr. H. Barrell and Dr. L. A. Sayce of interferometry in the National Physical Laboratory Standards and Light Divisions respectively.

An exhibition illustrating the papers presented and showing applications of interferometry, together with exhibits illustrating the work of the National Physical Laboratory in interferometry, is well and fully reported.

A very valuable feature of this publication is the account given of the discussion, rightly described as 'vigorous and progressive'. This is reported concisely, so inevitably omitting some of the amusing minutiae, such as the declaration by a specialist contributor that his entire spectroscopical experience had been obtained within the limits of the green line of mercury.

This book is a mine of information on the 'growing points' of interferometry. Published at a modest price, it is exceptionally good value for money and deserves a place in the library of all physical laboratories. J. DYSON

Mechanics, properties of matter, vibrating systems, elasticity

Mechanics and properties of matter. 2nd edn. By R. J. STEPHENSON. (London, New York: John Wiley & Sons, 1960.) Pp. x + 367. Price 60*s*.

The course in mechanics described in this book is given to students whose further studies may lie either in physics or in engineering. The treatment is almost entirely vectorial, Lagrange's equations being introduced but scarcely used, and at a level where there is a decided shortage of useful textbooks. A student working through this book, with its generous supply of examples at the ends of the chapters, will achieve a physical understanding of mechanics at present unusual, at least among British students, with no heavy strain on his mathematical technique.

The work includes an unusually complete treatment of forced harmonic motion with examples from several branches of physics, and leads up to a good discussion of rotating coordinate frames and the simpler properties of the gyrostat. Teachers looking for a book which will take their students to this standard without having the physics obscured by applied mathematics will find this most suitable.

It is a pity that less good can be said of the sections on properties of matter. These are on a level far below that of the mechanics, and their omission would make a cheaper and a rather better volume. Students using this book should be warned that the discussion of rolling wheels on page 45 (following an uninspiring description of some modern views on friction) must be extended to careful thought about the force systems at the axle before the matter can be said to be understood.

In spite of these criticisms, I consider this a very useful book indeed. Unfortunately, the price is rather high for students. G. WYLLIE

The mechanics of vibration. By R. E. D. BISHOP and D. C. JOHNSON. (London: Cambridge University Press, 1960.) Pp. xii + 592. Price 120*s*.

This latest of the Cambridge books on engineering, while wholly theoretical in character, aims at giving its reader a physical insight into the vibrational properties of simple and complex mechanical systems. This aim must be immediately admitted as fulfilled. The outlook of the book is generally academic, and attention is confined to the case of small oscillations, the topic of non-linear vibration having been reserved, along with those of instability, self-excitation, the vibration of rotating bodies and various further developments of matters here dealt with, for a second volume which the authors originally intended to write.

To the physicist, the book provides impressive confirmation of the genius of Lord Rayleigh, whose classic *Theory of Sound* forms the basis on which much of the theory is erected. The first three chapters deal with the behaviour of loss-free lumped-constant mechanical systems. The authors early introduce the concept of 'receptance', which is fundamental to the treatment throughout. Receptance at a point is the ratio of displacement amplitude to force amplitude, force being applied at the point in question or elsewhere. The receptances of simple mechanical elements and combinations of them are examined and reference tables given for common combinations.

Receptance is closely related to the mechanical admittance more commonly used by the physicist. To him, it will be a matter of slight regret that in the book there is no reference to nor use of the large body of directly analogous work existing on electrical network theory, if for no other reason than that impedance and admittance concepts were developed earlier in that field and are there much more fully documented. The authors have moreover eschewed the concept of mechanical circuits which completes the analogy with electricity, preferring when in doubt about the receptance structure to return to the differential equations for the network.

Chapter 2 introduces generalized coordinates and Lagrangian equations. These are then applied to many-degree-of-freedom systems and the idea of receptance extended to relate generalized forces and coordinates. Discussion of principal coordinates and orthogonality leads up to Rayleigh's principle.

The transition to distributed-constant systems is made using the stretched string as an example. The discussion covers generalized receptances in series form, characteristic functions, and the representation of arbitrary disturbances

in terms of them. Then the validity is examined of the unavoidable approximations introduced in idealizing real structures for theoretical purposes. Rayleigh's principle illuminates these considerations.

Chapters 6 and 7 apply these ideas to the torsional, extensional and flexural vibration of bars, making the usual engineering approximations regarding stress distributions over the cross sections, etc. A comprehensive treatment from the receptance viewpoint of such vibrations in single and composite elements is supplemented by valuable lists of receptances and tabulations of characteristic functions.

The next two chapters turn to the inclusion of damping, both viscous and hysteretic, in the equations of many-degree-of-freedom systems. They include discussion of how and when the principal modes cease to be independent, and of the coupling between them for small damping. The remaining two chapters consider non-steady vibration. A general technique is described for solution of free vibration problems with given initial conditions in complex structures, and finally excitation by arbitrary forces, using the Duhamel Integral, is discussed.

The mathematics used in the book is well within the range of undergraduate students, and extreme rigour is not sought. The text is lucid, orderly and very readable throughout. A special feature is the large number of examples (with answers) provided for solution, a set following each section of a chapter. These make the book especially useful to the teacher and advanced student, while to the physicist or engineer meeting problems of structural vibration it will be invaluable not only for its didactic qualities but for the grounding in analytical techniques which it provides and the many tabulations of useful functions and formulae which it includes.

The book, running to some 600 pages, is produced to the normal high standards of its publishers and is exceptionally free from errors.

G. G. PARFITT

An introduction to the theory of vibrating systems. By W. G. BICKLEY and A. TALBOT. (Oxford: Clarendon Press, 1960.) Pp. xiv + 238. Price 30*s*.

This volume is an extremely useful addition to the available literature dealing with vibration problems. The text may be considered in three parts: Chapters 1–5 which deal with the basic ideas of vibration in terms of the simplest mechanical and electrical systems; chapters 6–9 which develop the theory of Lagrange's equations and Rayleigh's principle and apply them to systems with several degrees of freedom; chapters 11–15, wherein wave propagation and vibrations in continuous systems are discussed; chapters 10 and 16 are on electro-mechanical analogies and non-linear systems respectively.

This planning works very satisfactorily in that energy concepts are introduced and explained in the second and third chapters and the reader is well prepared for their use in the work which follows Lagrange and Rayleigh. The other physical ideas which are important in the application of boundary conditions in continuous systems are also carefully explained. The similarity of the mathematical features in the formulation of vibration problems is emphasized throughout. The final chapter on non-linear systems provides an admirable introduction to a more specialized text like Stoker or Andronov and Chaikin.

There are exercises at the end of each chapter (except the last) which seem admirably chosen and whose answers as quoted are correct in the case of all samples checked.

For any student or research worker concerned with vibration problems who has the requisite mathematical background—little more than really thorough knowledge of an Advanced Level syllabus is needed—this book should prove invaluable, particularly since he should find the dire necessity for any other tutor infrequent.

The book comes from The Clarendon Press and is produced with the accuracy and quality which few publishers offer for the price.

M. J. P. MUSGRAVE

An introduction to applied anisotropic elasticity. By R. F. S. HEARMON. (London: Clarendon Press; Oxford University Press, 1961.) Pp. viii + 136. Price 35*s*.

Dr. Hearmon's monographs have become something of a *vade mecum* for anyone studying problems in classical elasticity, and it is most welcome to find many of the tables and formulae which originally appeared in these papers now collected together in the early chapters of the present work.

The author's approach is, to use the current jargon, 'phenomenological'—that is, he starts from Hooke's law and does not attempt to relate the elastic moduli to the concepts of atomic physics. The first three chapters establish the foundations of the subject, and Dr. Hearmon is at his most lucid in sorting out order from the chaotic jungle of contracted notations which has grown up. There is a brief description of methods of measurement of elastic moduli, but the emphasis is on the theory of the subject, in preparation for the later chapters of the book. On the whole the text is clear and concise, although the addition of a few diagrams would have made the section on crystal symmetry easier to follow. In the next two chapters the author considers a number of problems in elastostatics, and the remainder of the book is devoted to the analysis of dynamical problems, chapter six being mainly an account of some recent work on propagation of plane waves in anisotropic solids. Dr. Hearmon has devoted a considerable section of his book to this topic, and I think its intrinsic interest alone is sufficient to justify him in doing so.

The author states in his preface that he intends the book to be read by physicists and engineers rather than by pure mathematicians; this needs some qualification, as the treatment is predominantly theoretical and of necessity makes fairly heavy demands on the reader's mathematical ability. However, to any person seriously interested in the subject it is quite indispensable, and can be unhesitatingly recommended as a sound and perspicuous introduction. Production is of high quality; the type is clear and elegant, and equations are well set out. There is a useful list of principal symbols, and a comprehensive bibliography at the end of the book.

H. PURSEY

Large elastic deformations and non-linear continuum mechanics. By A. E. GREEN and J. E. ADKINS. (Oxford: Clarendon Press, Oxford University Press, 1960.) Pp. xiii + 348. Price 55*s*.

This book gives a connected account of the developments in large displacement elasticity over the last seven years. It will therefore be invaluable to all who work in this specialized field. To the much larger group of workers in the other branches of theoretical solid mechanics (of whom the reviewer is one) it will provide an enjoyable and stimulating account of a rapidly developing subject, even though they will probably have to take much of the analysis on trust. It is not

impossible that, with the next ten years, an understanding of this subject will become a prerequisite for research in solid mechanics.

Chapter 1 starts with the basic formulae of the subject (for proofs, reference is made to *Theoretical Elasticity* by Green and Zerna, Clarendon Press, 1954) and then the strain energy function is found for the basic crystal classes. This treatment may be of special interest to the crystallographer. Chapters 2, 3 and 4 give the few exact solutions of the finite theory, the plane strain theory and the plane stress and membrane theories respectively, and chapters 5 and 6 develop techniques of solution by successive approximation. Chapter 7 considers reinforcement of elastic materials by inextensible cords, which has application to tyres and to fire hose. Chapters 8 and 9 treat the stability and thermodynamics of the deformations. Chapter 10 considers the experimental determination of those physical properties of vulcanized rubber, which appear as arbitrary functions in the general theory. Chapter 11 is concerned with rheological equations of state and is therefore even more general than the rest of the book.

D. R. BLAND

Impact: the theory and physical behaviour of colliding solids. By W. GOLDSMITH. (London: Edward Arnold, 1960.) Pp. xv + 379. Price 90*s*.

There are seven chapters: Introduction (3 pp.), Stereomechanical impact (16 pp.), Vibrational aspects of impact (52 pp.), Contact phenomena produced by the impact of elastic bodies (55 pp.), Dynamic processes involving plastic strains (95 pp.), Results of impact experiments (47 pp.), and Dynamic properties of materials (68 pp.). This last chapter is relatively short and perhaps a little biased towards the properties of metals. But much indirect information on experiments is interspersed with the theory in the earlier parts. The bibliography is extensive. Besides a main list of 442 alphabetically ordered references, there are substantial collections of supplementary references to the individual chapters.

The book covers a very wide range of topics in applied mathematics and applied physics. Inevitably, the information is highly condensed. Also inevitably, readers interested in some special topic will be able to think of references that have been omitted. For instance, H. Kolsky's work on pulse propagation in idealized polymers might have deserved a place in section 3.10. Generally speaking, the way in which reference numbers are given as superscripts to words of the text, without other indication of their relevance, makes for hard work on the reader's part.

Having thereby hinted at some possible criticisms, one must add quickly and emphatically that in many ways the book is admirable. Physical principles and mathematical arguments are presented clearly and accurately, despite the necessary brevity. The way in which Professor Goldsmith can short-cut several paragraphs of explanatory text is exemplified by the idealized stress–strain curves of Figure 128, page 207. For an example of lucid mathematical exposition see section 2.8 on the Lagrangian equations for impulsive motion.

There is no doubt that this work will prove extremely useful to any physicist who has a strong interest in applied mechanics. Beyond its value as a source of reference, it gives the research worker a good demonstration of the application of the theory to complex phenomena. The treatment of impact problems requires a combination of classical mechanics and elasticity theory with plasticity theory, and, occasionally, with hydrodynamics. For success, the choice of the right idealizations is all-important. This requirement is, of course, central to theoretical physics in general. Hence the book can teach not only about impact, but also about the methodology of applied mathematics.

F. C. ROESLER

Mechanical waveguides. By M. R. REDWOOD. (Oxford: Pergamon Press, 1960.) Pp. ix + 250. Price 50*s*.

The subject-matter of this book is more precisely conveyed by the sub-title 'The propagation of acoustic and ultrasonic waves in fluids and solids with boundaries'.

Starting from definitions of strain to first order, stress and the generalized Hooke's Law, the equations of motion for solids and fluids are derived. Attention is then immediately focused on the wave equation and solutions appropriate to unbounded isotropic media are developed. These solutions are used to solve reflection–refraction problems at the plane interface between two semi-infinite media. Succeeding chapters (3, 5, 6, 7) deal with solutions for continuous progressive waves in plates and circular cylinders and Chapter 8 gives a brief résumé of solutions for cylindrical rods of other cross sections and for circularly cylindrical shells.

Most, if not all, of this material may be found in Love's *Mathematical Theory of Elasticity* or other well-known texts as indeed the author implies in his Introduction. However, it is here very competently collated and is of immediate relevance to the other large fraction of the book, Chapters 4, 9, 10, which concerns the propagation of pulses in fluid and solid waveguides, viz. along plates and circular cylinders.

In this field, the author is himself a pioneer and he has given us between the covers of this book an exposition of much important work which was hitherto widely scattered.

The final chapters 11, 12, 13 on multi-layer problems, resonance and anisotropic media are essays of introduction to the respective subjects. There are five useful appendices.

The references given at the end of each chapter are extensive; those particularly relevant to the text are itemized separately.

For the electrical engineer or the physicist with expertise in electronics who starts to face problems in fluid and solid mechanics, this book should be of the greatest value since while dealing largely with mathematical theory, Dr. Redwood never fails to remind him of the difficulties and assumptions involved in relating theory and experiment. The more mathematically inclined may wish he would preserve cyclic order (e.g. Eqns (1.3), (1.4), p. 4, or pp. 260, 266) but they will find much to appreciate.

There remain a few misprints, mostly suffixes or affixes (none noticed by the reviewer gives rise to serious ambiguity) and the question of layout and readability of the mathematical expressions (see Eqn (1.7) or p. 149). Ideally, a larger size of page would be chosen for such a text and it is to be hoped a more spacious edition may be considered in the future.

M. J. P. MUSGRAVE

Non-destructive testing, ultrasonic devices, X-ray microscopy, etc.

Techniques of non-destructive testing. Ed. C. A. HOGARTH and J. BLITZ. (London: Butterworths, 1961.) Pp. vii + 216. Price 40*s*.

This volume contains the substance of a course of lectures at Brunel College of Technology together with some additional material. It is intended that it should be of use to "inspectors

and others concerned with the application of non-destructive testing techniques in industry and commerce". (One author uses the description "non-destructive test technicians".) This intention has determined the general choice of subjects and level of treatment.

Radiology is covered by a chapter on general principles followed by a chapter on the methods which are in use. There are two parallel chapters on Ultrasonic testing and single chapters on Radio-isotopes, Magnetic methods, Eddy current methods, Penetrant methods, and the Thermal comparator (a very useful chapter on a technique which is not widely known). A final chapter gives an account of the use of non-destructive testing methods in an organization mainly concerned with the inspection of aircraft components.

The sections describing basic principles are elementary and generally adequate. The illustrations are simple and direct except in one case where, as it stands, the diagram is not sufficiently indicative (Fig. 5.1). In the sections dealing with techniques there is much material of direct concern to the practical man (technician perhaps) who is using the methods. It is not by any means certain that anyone could use the book as a complete guide to the application of any of these methods although for some of them a considerable amount of detail is given.

The book will evidently serve as an introductory text for people who wish to become inspectors or technicians in this field. It will help those who wish to get a broad view of the subject. It is not comprehensive enough to enjoy the full status of a complete practical treatise. Physicists and others who are concerned with the physical principles and with properties of materials which could form the basis of new tests will, quite obviously, not use this book. In so far as it explains the physical principles of existing techniques it may however be considered to represent an attempt at communication—necessary and desirable "all along the line".

K. W. ANDREWS

Ultrasonics. By B. CARLIN. 2nd edn. (London: McGraw Hill, 1960.) Pp. ix + 309. Price 89*s*.

This is the second edition of a book by an American engineer. The author is mainly concerned with the practical design and application of ultrasonic devices, rather than with the theoretical background to the subject. This approach is particularly noticeable in the first chapter, which is intended as a general introduction to the subject. The manner in which this introduction is presented could confuse a newcomer to the subject, or irritate anyone more familiar with the concepts summarized. In particular, the brief discussion of wave motion is unfortunate, and the descriptions of a 'wave-train' and a 'typical' pulse shape are misleading.

The second and third chapters are concerned with piezoelectric materials and their use in ultrasonic transducer assemblies. A useful chart for determining the electromechanical characteristics of various piezoelectric crystal slices is reproduced. The properties of the major cuts of quartz crystal slices are discussed at some length. To have summarized these concisely and carefully might have been more valuable. The third chapter includes some figures illustrating the construction of a variety of crystal mountings. Some of these (e.g. Figs 3–20) are neatly drawn. Others (such as Figs 3–21) are apparently reproduced from other publications and include symbols and subsidiary sketches which are not referred to in the text.

The chapter devoted to magnetostriction includes descriptions of most of the transducing arrangements and techniques in general use. The occasional use of unexpected abbreviations is distracting. For example, 'the Joule' is used when 'the Joule effect' would have been more accurate. Magnetostrictive delay lines are mentioned in this chapter. The author states that the use of these "is not so common as that of the crystal types". This comparison of magnetostrictive and piezoelectric delay lines may not now be as valid as at the time of the 1949 first edition, when computer applications of the magnetostrictive delay line would have been much less than in more recent years.

The author includes an enthusiastic chapter devoted to miscellaneous ultrasonic transducers such as whistles, sirens and various electromagnetic devices. Consideration is then given to certain measurement techniques applicable to ultrasonic experimentation. A table of 'typical' exposure times is given for photographing cathode-ray tube traces, together with a suggested method for mounting a camera for this type of work. The text confirms that the exposure table refers only to a particular phosphor, although no information is given on the cathode-ray tube beam current and final anode potential for which these results were obtained. Unfortunately many of the otherwise useful cathode-ray tube trace photographs actually included in the book exhibit distortions attributable to poor photography.

A chapter on high-power ultrasonics and two on non-destructive testing include considerable reference to practical circuit arrangements and actual instruments and equipment. In an interesting chapter on ultrasonic effects a large number of different phenomena are described. Further ultrasonic applications such as materials testing, medical techniques, cleaning, soldering and drilling are broadly surveyed in a final chapter.

Some small printing errors exist, such as the use of the Roman XI for a chapter heading (page ix), the spelling of text as 'test' (page 31, line 18), and the omission of the acute accent in 'Lamé's constant' (page 31, line 23). The book includes numerous references, many of which will interest the reader seeking to widen his knowledge of practical work carried out in recent years. An adequate index is provided.

M. A. SNELLING

X-ray microscopy. Cambridge monographs on physics. By V. E. COSSLETT and W. C. NIXON. (Cambridge: University Press, 1960.) Pp. xiv + 406. Price 80*s*.

Radiography deals with the x-ray shadows of opaque objects at unit magnification and their observation without visual aids. X-ray microscopy is interested in finer detail and looks either at shadows which are geometrically, by point projection, enlarged (projection microscopy) or at well defined shadows under a microscope (contact microradiography) or it uses a combination of both methods. X-rays of wavelengths in the 10 A region are normally employed, and by both techniques resolution down to about 0·1 microns is now being achieved. Attempts at using optical imagery proper with x-rays have not so far been very successful. On the other hand, indirect methods of using locally emitted x-rays to obtain information about the composition of surfaces (x-ray emission microscopy) have made a promising start. The last decade has seen a remarkable increase in the interest in x-ray microscopy which has led to considerable progress. Members of the growing community of x-ray microscopists and others interested in the field will find in Cosslett and Nixon's text a review of the present position of the art, well documented by a systematic collection of references.

The book reads easily. Even non-physicists should be able to read it without recourse to other sources of knowledge as the basic facts and some simple calculations are explained at considerable length. The book contains also a large variety of detailed information. On the other hand, the data provided are not sufficient to enable a newcomer to set up his own apparatus. Such detail as is given certainly enlivens the text, but on occasions the reader might be puzzled by its selection. It seems curious to find statements such as "The penetration of x-rays is high compared with visible light because of their much shorter wavelength . . . ", or to find winding data given for a magnetic lens without corresponding detail about the iron core and the pole pieces. Again, a large number of reproductions of beautiful microradiographs illustrate various applications of the art, but text and captions tell the reader only what they are and how they were obtained, and often do not indicate their significance or what useful or new facts the photographs reveal.

The active x-ray microscopist will enjoy reading the book as a survey. It will widen his horizon and make him review the scope of his work. The prospective x-ray microscopist can from the examples form an opinion about the applicability of the various techniques to his problems and he will get an idea of the experimental difficulties he may encounter. He will not find a blueprint for his work, but will acquire a sound background for further enquiries.

W. EHRENBERG

Progress in semiconductors. Vol. 5. Edited by A. F. GIBSON, F. A. KRÖGER and R. E. BURGESS. (London: Heywood and Company, 1960.) Pp. viii + 318. Price 63*s*.

The book maintains the high standard set by preceding volumes in the same series. A large part of it is devoted to optical properties. T. P. McLean deals with the structure of the absorption edge spectrum on the basis of direct and indirect transitions and also gives a brief description of experimental techniques. By the use of gratings in place of prisms, the resolution of infra-red measurements in the neighbourhood of 1 micron has been increased by about one order of magnitude. The effect of exciton generation is considered and it is shown how absorption measurements can lead to information not only about the electronic states near the band edges but also about the vibrational spectrum of the semiconductor. This article is suitable for newcomers to the subject.

T. S. Moss contributes a concise review of indium antimonide, dealing with the various absorption mechanisms and also with recombination, magneto-optic effects, photo-effects, conduction properties and, quite briefly, with several applications. A wide field is covered, but the references are very good and make the review, amongst other things, a useful guide to the literature.

The subject of magneto-optical phenomena is further developed by B. Lax and S. Zwerdling who deal specifically with magneto-absorption, the Zeeman effect of excitons, the Faraday rotation and magneto-plasma effects. There is an interesting discussion of the manner in which the measurement of these effects supplements and extends the information derived from cyclotron resonance. The authors indicate how the combination of high resolution spectroscopy, low temperatures and very high magnetic fields could be further exploited, not only for research on semiconductors but for other solids as well.

A review of recent researches on graphite has long been overdue. R. R. Haering and S. Mrozowski discuss the band structure and electronic properties, including such topics as the oscillatory Hall effect. J. Appel gives an account of recent work on heat conduction in semiconductors from phenomenological and fundamental points of view. Again, not suitable for newcomers. A discussion of chemical binding by E. Mooser and W. B. Pearson covers a wide range of substances and should be particularly useful in the exploration of new materials. It includes a new and interesting discussion of the relationship between surface structure and surface states. Semiconductor surfaces are also the subject of a separate article by T. B. Watkins. This is a fine introduction and deals with surface conduction and recombination as well as the corresponding field effects. There are ninety-nine well chosen references.

Review books of this kind can provide their most valuable service by offering readers an opportunity for acquainting themselves with recent progress in fields other than their own. To do this well, the various articles should be designed essentially to introduce, to survey, to stimulate and to teach. It is not always necessary to go to the limits of theoretical complexity and sophistication and, indeed, many reviews are the better for avoiding this. From this point of view, the contributions to this volume could be more uniformly effective than they are, but the book remains a valuable aid to the semiconductor physicist and can be safely recommended.

H. K. HENISCH

Nuclear physics

Progress in nuclear energy, series IV. Technology, engineering and safety. Vol. III. Edited by C. M. NICHOLLS. (Oxford: Pergamon Press, 1960.) Pp. viii + 448 + xii. Price £5 5*s*.

This is the third volume in the section of 'Progress in Nuclear Energy' dealing with technology, engineering and safety. It is, of course, impossible to cover such an all-embracing field in a representative manner in a single volume, and the selection of papers therefore appears somewhat arbitrary. The intention is to review selected topics in the field, rather than to publish original contributions; the usefulness of the book therefore depends upon the lasting value of such reviews.

Two papers in the engineering section deal with pumps used in the nuclear engineering industry. The first is a fairly comprehensive review of the performance of gas bearings in gas compressors, and should provide a useful introduction to anyone wishing to design such a compressor. The second is little more than a catalogue of the various designs of mechanical and electromagnetic pumps which have been developed for handling difficult liquids. The discussion of their relative merits is rather superficial, and would not enable the design engineer to make a firm choice between them.

The technology section includes a useful review of steam as a reactor coolant, and four papers dealing with the chemistry and metallurgy of fuel and fuel elements. One of these sets out the solutions to a range of steady and transient thermal stress problems, and will certainly be of use to the designer of metallic fuel elements. One feels, however, that the usefulness of the paper would have been enhanced by more frequent comparisons with experiment.

The final two sections deal with criticality and with reactor safety problems. The first paper presents a very comprehensive review of eighteen criticality accidents which occurred in the United States in the period 1945 to 1956, together with a most interesting analysis of each. One is left with the impression that a great deal of care and effort has been devoted to the analysis of data which, though sometimes

rather sparse, are of enormous value. Since such accidents do unfortunately occur, it is perhaps as well to be prepared, and the second paper considers the steps which should be taken to control such emergencies. It is noteworthy that in a number of cases the serious consequences of an accident have been materially reduced when the people concerned have been adequately trained beforehand. It is to be hoped that as much effort is normally devoted to the analysis of a process before it is put into operation, and the remaining papers in this section should be of use in this respect. Two are devoted to the criticality of homogeneous solutions and mixtures, the first dealing with uranium–water systems, and the second with plutonium–water systems. In the case of uranium it has been possible to test the theory against a considerable volume of experimental work; in the case of plutonium, experimental data are rather sparse. The remaining papers consider neutron interaction in an array of assemblies of fissile material, and mathematical models used in assessing criticality problems.

Finally, under the heading of Reactor Safety there are papers on stored energy in graphite, control of effluent from reactors and chemical plant, and reactor emergency shut-down equipment. The latter two are unlikely to be of general interest, but the paper on stored energy can be recommended as a clear and interesting summary of experimental results and their interpretation.

On the whole this collection of reviews is rather more successful than most of its type, mainly due to the useful and comprehensive collection of papers on criticality problems.

W. B. HALL

Kernenergie Technik. By H. ENGEL and K. O. THIELMANN. (Munich: Verlag Moderne Industrie, 1960.) Pp. vii + 300. Price 68*s*.

The authors of this little book have attempted the formidable task of reviewing nuclear energy practice in all its aspects within the compass of only 300 pages. It opens by describing all the basic techniques of the atomic energy industry, and this section covers physics, engineering, metallurgy and chemistry, both as applied to the building and running of reactors and as relevant to the production of the materials used. From this point it proceeds to deal with the particular types of reactor at present constructed, the economic implications of nuclear energy and its uses, and concludes with a chapter on the outlook for a practical source of thermonuclear power.

The overriding impression gained on reading the book is one of drastic condensation dictated by the attempt to be exhaustive, and reviewing the result one cannot help doubting the wisdom of making such a short volume so all-embracing. In a book apparently written mainly for engineers it is understandable that nuclear physics is dismissed in 23 pages, but not that reactor physics should only deserve a further 24, in which diffusion theory is not mentioned and the spatial distribution of neutron density only incidentally touched on. Even more surprising in the circumstances is the restriction of heat transfer problems, which, from the engineer's point of view, are the key to future reactor design, to only 16 pages. Had the authors, for example, attempted a little less by confining themselves to technical aspects as at present known, leaving economic questions and future developments to another volume, then the space thus gained would have been most useful. Probably the most serious consequence of this brevity is that only the reader who is already expert in this field will realize quite how much that is fundamental has been omitted.

Having once accepted the aim of the book, a most attractive feature is the inclusion of many clear graphs giving immediate numerical values of most quantities involved in reactor design. Indeed, here lies its greatest value to the practising reactor technician and it should prove invaluable for obtaining figures to put into all kinds of 'back of envelope' calculations regarding reactor operation and design. The reader is helped in this by the provision of worked examples to each chapter, though again there is a danger that the unwary may not realize just how much these have been simplified. Although errors and omissions in the individual sections can be found, the authors are to be very much commended on their selection of the items to be included, which must have been a truly difficult task. It is therefore surprising that, in the section on material production, he has chosen to make no mention of the chemical exchange method of deuterium production. The book concludes with a useful although short bibliography.

To sum up, for anyone wanting a short overall review of the atomic energy field, and to the expert wanting to obtain some approximate figures quickly, this book, read with an understanding of its limitations, will prove most useful.

M. J. POOLE

Fast neutron physics: Part I. Techniques. Interscience monographs and texts in physics and astronomy, Vol. IV. Edited by J. B. MARION and J. L. FOWLER. (London, New York: Interscience Publishers, 1960.) Pp. xiv + 983. Price 218*s*.

Most nuclear physicists are aware that the literature of slow neutron physics, because of its connection with the nuclear reactor, is extensive and detailed. However, it may surprise many that the physics of fast neutrons now presents so many diverse features that a volume of nearly 1000 pages has been produced to describe merely the instrumental side of the subject. The theory is to follow in Part II, which may well be of comparable size. The reason for this scale of effort is partly that the subject is treated at full reference-book level, and partly that nuclear physics presents such a closely interlinked pattern that a discussion of one branch inevitably brings in references to many others. In the present encyclopaedic book the editors have defined fast neutrons to be those with energies lying between 1 kev and 'several hundred' MEV, although the emphasis in the 24 articles is mainly on energies below 40 MEV. Each article is meant to be complete in itself and this leads to a certain amount of unavoidable duplication, probably noticeable only when the book is read through, rather than consulted.

The first section, Neutron Sources, deals with radioactive sources as well as the monoenergetic type of source based principally on the (p, n) and (d, n) reactions. Much of the remainder of the book is concerned with measurements made using the $^{3}H(d, n)$ and ^{7}Li (p, n) reactions and the latter receives detailed treatment in a chapter of 43 pages, 17 of which are tabular. Section II contains 5 chapters on Recoil Detection Methods in which the applications of gas-filled counters, scintillators, and visual detectors are discussed. The volume was apparently completed too soon to include more than a very brief mention of the new solid-state detectors, but it is surprising that Čerenkov counters for high energy neutrons and semi-automatic methods of track analysis for photographic plates and bubble chambers are

not mentioned. Section III, Detection by Neutron-Induced Methods, covers the popular long counter, the liquid scintillator, the gaseous scintillator, the helium-3 spectrometer and fission detectors. The final section, on Special Techniques and Problems, includes monographs on time-of-flight techniques, neutron flux measurements, shielding problems and radiobiology. The last is an excellent elementary survey of the biological effects of radiation in general. The final chapter gives a brief and welcome account of some computer techniques.

The articles in this book are authoritative, detailed and informative; they generally cover work which was available at the end of 1958 or early in 1959. The 33 authors, most of whom are attached to large research laboratories, have spared no pains to illuminate all facets of their subjects. As a result, the general reader will find the detail excessive at many points, but this is what is required of a reference book, whose price will in any case probably confine it to libraries. The book is well produced and will be of service to all who are concerned with the design of fast neutron experiments; it is to be hoped that its companion volume will be of equal use to those who wish to interpret their observations.

W. E. BURCHAM

Controlled thermonuclear reactions—an introduction to theory and experiment. By S. GLASSTONE and R. H. LOVBERG. (London: Van Nostrand, 1960.) Pp. xvi + 523. Price 42*s*.

This account of the present state of controlled thermonuclear research may possibly be the last of its kind. It contains a description of a number of possible devices which it had been hoped might be developed into thermonuclear reactors, together with accounts of the experiments inspired by these hopes. In part, it represents a fuller, more detailed and more technical revision of Bishop's history of the American fusion program 'Project Sherwood', although reports of British, European and Russian work are included.

The picture presented is, superficially, rather discouraging; the principal result of several years' intensive research, including much elegant theory and a number of ingenious experiments, has been the discovery of several unexpected difficulties in the path toward a controlled thermonuclear reactor. This stresses what has long been conceded intellectually by workers in the field, but only much more recently emotionally felt—that attainment of controlled thermonuclear reactions will demand a complete understanding of the physics of hot plasmas, and is unlikely to be in the grasp of some ingenious and fortunate gadget. Such a conclusion is not without its consolations to the physicist since plasma physics is a most exciting and undeveloped branch of classical physics. Besides, while not producing too facile an optimism, recent experiments which have involved the compression of plasma in magnetic mirrors have given encouraging indications of heating and confinement; none the less, I expect that those books on the subject to appear in the near future will bear titles less suggestive of the inspiration of the research and will be organized about topics in the physics of plasmas rather than about speculative thermonuclear reactors.

Although the present work appears to have been assembled rather than written, it is refreshing to encounter a survey from the experimental viewpoint and one which lays more stress on measured numbers than on hopeful speculation or elegant formalism. Even if the present dearth of books on plasma physics were absent, this work could be recommended as an introductory survey to interested physicists, and as required reading for plasma experimenters.

W. B. THOMPSON

Materials for nuclear engineers. Edited by A. B. MCINTOSH and T. J. HEAL. (London: Temple Press, 1960.) Pp. ix + 373. Price 75*s*.

The editors have set out to gather the latest information on the properties of a range of materials used by nuclear engineers, and to present the information in a form suitable for engineers. Individual chapters have been written by members of the United Kingdom Atomic Energy Authority who have been engaged in the work concerned, and the range of the materials selected is therefore closely connected with the United Kingdom's nuclear energy programme.

There are chapters on: materials information—its uses in nuclear engineering, uranium, plutonium, thorium, ceramic fuels, graphite, magnesium, beryllium and zirconium. Each chapter is broken down into sections. The main ones of which are, the nuclear and physical properties of the material, its mechanical properties, its compatibility with other materials, its behaviour under irradiation, its alloying properties and fabrication. At the end of each chapter all the information on a particular material is summarized in an appendix, and a comprehensive list of references given.

The book is well laid out and is a useful accumulation of technical data. Little attempt has been made to correlate the results with theory. The preferred aim of the book has been towards a deliberate omission of much of the metallurgical discussion and the significance of the information to the engineer discussed instead in the hope that he may derive 'a feel for the characteristics of the materials' he uses. In this, the book to a certain degree fails. Not enough effort is made to give a physical interpretation of the processes involved, and the main use of the book to an engineer would be as a handbook of information.

Probably due to the large amount of data available, the chapter on uranium is too condensed. As important a property to the engineer as the coefficient of thermal conductivity is not discussed and only the results from two unpublished papers quoted without comment in the appendix. The chapter on beryllium is the only one which omits any information on the behaviour of the material under irradiation.

Although ten authors have contributed, a fairly consistent presentation is achieved in spite of this inherent difficulty of a composite volume. The range of the materials selected is fairly comprehensive. Further chapters on control rod and vessel materials would have been useful. J. D. YOUNG

Nuclear propulsion. Edited by M. W. THRING. (London: Butterworths, 1960.) Pp. 300. Price 50*s*.

The book deals with the problems associated with nuclear marine propulsion and nuclear-powered aircraft and rockets. It can be divided into two sections. In the first half basic nuclear and reactor theory is first considered, followed by a chapter on the thermodynamics of jet and rocket propulsion. This is then followed by chapters on the various aspects of reactors—design, control and instrumentation, and heat transfer. The second half of the book is principally concerned with nuclear propulsion and its possibilities for ships and rockets. Some interesting side lines, such as the handling of working fluids and aspects of life in sealed cabins are also covered.

As is to be expected in a book that covers a variety of specialized topics, the contents consist of a number of articles written by different authors, which technique can lead to some repetition and to a sense of diffuseness about the general purpose. This book does not avoid these pitfalls and it is difficult at the end really to pick out some of the problems associated with nuclear propulsion or the manner in which they have been, or might be solved. Also the notation used is not always consistent throughout.

However, many sections of the book make interesting reading and an appreciation of the usefulness of nuclear reactors for propulsion is certainly obtained. Some topics could well have been expanded, for instance shielding, which is not accorded a chapter, and heat transfer, and the uncertainties in the calculations and their effect of the fuel economics pointed out. This also applies to the uncertainties in the masses of nuclear fuel required and its expected lifetime. No indication is given on the book's paper cover of the type of people for whom it was written. It is probably most useful for graduates who are considering entering the nuclear reactor field and wish to ascertain the possibilities of nuclear propulsion. V. S. CROCKER

Eléments de physique nucléaire. By D. BLANC and G. AMBROSINO. (Paris: Masson et Cie., 1960.) Pp. 238. Price 30 NF.

This book aims to provide engineers and others coming into atomic energy for the first time with the scientific background they need to carry on with their work. It is in no sense an academic textbook of nuclear physics: theoretical matters are mostly avoided, and formulae are often quoted in the style of a handbook with only a brief accompanying text.

Following on a review of elementary atomic structure and radioactivity, there comes a detailed account of the dynamics of nuclear reactions and the concept of reaction cross section. From a brief survey of nuclear reactions in general we turn quickly to a fuller account of neutron physics, fission and chain reactions. The point of view is always that of the nuclear physicist; accordingly, we learn in detail about the neutron balance in a reactor, but not at all about problems of heat transfer or chemical processing.

The next chapter deals with nuclear structure and nuclear forces. Here more sophisticated matters are discussed—nuclear models, statistics, parity, exchange forces, mesonic atoms, and strange particles—but the treatment is so brief that the uninitiated would surely find it puzzling. The final chapter reverts to more practical topics, and gives a useful summary of the thermonuclear problem, with some drawings and photographs of experimental apparatus.

In its choice of topics the book is thus a rather curious mixture: it reflects perhaps the interests of a 'pure' nuclear physicist thrust into a milieu where his subject has become 'applied'. As such it could be useful for broadening the mind of an undergraduate or graduate physics student. As an exposition of the physics of atomic energy for engineers, on the other hand, it is much too uneven in scope and difficulty. Without a doubt the English-speaking reader should turn to Glasstone instead. H. R. ALLAN

Radioisotope laboratory techniques. 2nd edn. By R. A. FAIRES and B. H. PARKS. (London: George Newnes, 1960.) Pp. xi + 224. Price 25*s*.

This manual of practical advice and information has, since its original appearance in 1958, become recognized as an indispensable aid in the design and operation of laboratories using radioactive materials. The new edition has been revised in detail and the chapter on health physics has been rewritten in the light of current international recommendations on radiological protection. J. M. A. LENIHAN

Radioastronomy

The radio noise spectrum. Edited by D. H. MENZEL. (Massachusetts: Harvard University Press; London: Oxford University Press, 1960.) Pp. viii + 183. Price 60*s*.

The photograph of a steerable radio telescope on the dust cover of this book will warn the communications engineer that the contents may not be quite what he would expect from the title. The book is, in fact, based on papers read at a conference on Radio Noise held at Harvard Observatory in 1958, and deals with 'noise' as understood by Radio Astronomers. A glance at the contents, however, shows that even for Radio Astronomers the title is not very precise. There are, for example, chapters, each written by a different author, on such non-noise-like subjects as 'The aurora and radio-wave propagation' (A. M. Peterson), 'Ionospheric scintillation of radio waves of extraterrestrial origin' (R. S. Lawrence), 'Meteor Scatter' (V. R. Eshleman), 'Interstellar Hydrogen' (T. Gold).

It is natural in a book based on a conference that much of the material should be a summary of published papers. The most complete of these are Peterson's on the aurora and Eshleman's on meteor scatter. Aaron's chapter on very low frequency noise and the allied subject of magnetic fluctuations with a period of order one second is a useful survey of a little-known subject.

Amongst the material which has not previously been widely published are a chapter in which Gold puts forward the suggestion that 'Whistler-like' noise might result from the dispersion of impulses which arise in the sun and travel to the earth along electron streams, and a chapter in which Hawkins describes an unsuccessful search for radio noise originating in meteors.

One of the most useful chapters is that by the Editor, Dr. Menzel, in which he surveys what is known about radio sources in the sky, and summarizes the knowledge in a series of contours from which the noise intensity in any direction and on any frequency can be determined. This is probably the only chapter which will interest the communications engineer, who might be misled by the title. The Radio Astronomer, on the contrary, might ignore the book because of its title: if he did he would miss much of interest.

J. A. RATCLIFFE

Theoretical physics

Small particle statistics, 2nd edn. By G. HERDAN. (London, Toronto, Sydney: Butterworths, 1960.) Pp. xxiii + 418. Price 80*s*.

Certainly in studying particulate systems some knowledge of statistical methods is required for the design of experiments and the critical appraisal of techniques as well as analysis of observational data. Much of the necessary information is to be found in Dr. Herdan's book, and to many workers the first edition is already a useful source of reference.

That a statistician like Dr. Herdan should enter this field with such enthusiasm to 'establish the statistics of particles . . . as a branch of statistics, theory and practice, in its own right' is to be welcomed, but he is confronted with difficulties. The science of small particles is essentially a

practical one. Its statistical aspects must be considered against a background of physical theory and experimentation and in these matters Dr. Herdan seems to be less at home. One gets the impression that much of his book has been written around a card index of literature abstracts. Although these have not always been thoroughly digested, the many references are a valuable feature.

The new edition follows a similar pattern to the old. The eight chapters of Part I deal with fundamental concepts of frequency distributions, units and principles of measurement, distributions and averages, sampling procedures, standard forms of distribution function, graphical representation and differences between determination results. They contain new sections on automatic counting and sizing, transformation between distributions, the Rosin–Rammler law and the analysis of variance. Part II comprises four chapters on relations between variable characteristics, correlation methods, particle size and other characteristics, and physical and chemical properties; there is only a little that is new. Part III has three chapters on the attainment of specified fineness, industrial mixing and inhomogeneity of polymers. The latter two contain much new material. In Part IV the two chapters on principles of statistical theory and the design of experiments remain much the same.

Part V, as before, has different authorship. In the present edition, Dr. M. L. Smith has been joined by Mr. W. H. Hardwick and Mr. P. Connor and they have enlarged the section to nine chapters. Here, under the title Experimental Design and Experimental Errors of Particle Size Determination, is hidden an excellent succinct review of the leading methods of particle size analysis together with a well-balanced appraisal of their particular applications and limitations. The ground covered includes optical and electron microscopy, automatic counting and sizing (new), sedimentation, centrifugal sedimentation, radiometric techniques (new), dispersions in gases (new), and the determination of surface area by permeability, by photo-extinction and by adsorption.

One can find many points on which to criticize this volume; in giving a few I do not wish to question its general usefulness. Table 2.4 presents sedimentation analysis data extending down to a particle diameter of 0·032 μ yet Table 3.2 gives the lower limit of the method as 1 μ–2 μ. In the sections on automatic counting and sizing, features of 'spot' scanning are erroneously attributed also to 'slit' scans (p. 49); it is not true that single spot scans cannot be applied when the number of interceptions exceeds the number of (circular) particles (p. 339) or that dual beam cathode-ray tubes are used for double spot scans (the two beams could not be isolated by the detectors) (p. 341). Figures 5.3 to 5.6 give he effects of sieve loading on size analysis results but the size of sieve on which the load is placed is not stated. Dust particles smaller than 5 μ (other than soluble ones) do not pass from the lungs into the blood (p. 224). Weight distributions seem to puzzle the author, who introduces phase space and eigenvalues to aid the discussion (p. 291)! Stokes's diameter is incorrectly defined (p. 299) as the diameter of a sphere of the same surface. In electron microscopy one does not photograph the image on the fluorescent screen but records directly from the electron beam on to the photographic plate (p. 336).

A notable omission from the statistical sections is any reference to the problem of particle overlap in randomly deposited samples evaluated by microscopy.

Despite its limitations, the new edition, like its predecessor, will be a valuable source of information on statistical matters connected with particle size analysis. I am sure that many practical workers, after reading the last Part, would also welcome a full volume on methods of particle size analysis by its authors—perhaps with a concise appendix on the requisite statistics.

W. H. WALTON

Field computations in engineering and physics. By A. THOM, C. J. APELT and G. F. J. TEMPLE. (London: Van Nostrand, 1961.) Pp. vii + 165. Price 30*s*.

This book summarizes the work of Professor Thom and his students and colleagues spread over nearly forty years, on the numerical solution of partial differential equations of elliptic type, with practical applications in hydrodynamics. Fourteen short chapters discuss the derivation of finite-difference formulae, in one and two variables, the method of 'squaring' for solving the resulting algebraic equations relevant to elliptic problems, and its application in various contexts to the solution of the equations of Laplace and Poisson in two dimensions, problems with 'free' boundaries, the biharmonic equation treated as simultaneous second-order equations, the Navier–Stokes equation, compressible flow, and axially symmetric problems. The whole is illustrated with numerical examples of a non-trivial nature.

The method of 'squaring' is described in detail, and used with a variety of formulae, many of which are not widely known, for giving better accuracy and quicker convergence. A feature of the book, particularly in connection with Laplace's equation, is the use of the stream function and potential function either as dependent or as independent variables, with the aim of avoiding curved boundaries, and also of other conjugate functions, such as direction and a function of the velocity, for simplifying the treatment in flow problems. Special integration formulae are developed to represent the Cauchy–Riemann equations. The functions of velocity and direction are particularly valuable for the production of 'symmetro-morphic figures' which consist of an integral number of 'squares' in the new plane, and which approximate to specified dimensions in the original plane. We are then solving a problem in a field of slightly different dimensions from the original, but the solution is made easier by the avoidance of irregular stars, and the slight difference in the two shapes is not important for many problems.

The chapter on convergence and propagation of error uses fairly elementary methods to find what other authors call 'the largest eigenvalue of the iteration matrix', and the results agree with those of more rigorous mathematical investigations for the simple 4-point formula. Other results are given for the standard 9-point formula (the '20' formula) and for another which uses a wider spread of points (the '100' formula). The more elaborate formulae give more rapid convergence and the relevant matrices have smaller inverses, so that what a 'relaxer' would call the 'residuals still unliquidated' have a smaller effect on the required values of the function.

The treatment of singularities of various kinds is also discussed, usually on an *ad hoc* basis and often lacking rigorous mathematical justification, but with methods which apparently work. This, in fact, is the value of this book. The methods work, and they represent the artistic branch of numerical computation to which engineers and other scientists, of whom Professor Thom is an outstanding example, have made such valuable contributions. But not all the methods have been programmed for the automatic machine, and some still need human judgment which is

difficult to code. It is a challenge to the mathematician and numerical analyst both to 'mechanize' these processes and also to provide the rigour needed to transform the *art* of computation into the *science* of computation. L. Fox

Plasma physics. By J. E. Drummond. (London: McGraw-Hill, 1961.) Pp. xiv + 386. Price 97*s*.

Plasma physics by J. E. Drummond is a collection of articles on plasma physics, by various workers in the field, which should prove to be very useful to all those who work in the field of electron gases, as well as to those who are interested in the many-body problem more generally.

The book begins with a brief introduction to the whole subject, in which the history is summarized and some of the main concepts of plasma theory, both classical and quantum mechanical, are described. This section has an extensive bibliography, which should be helpful, especially to those who are new to the field.

Following this introduction a considerable range of topics is treated, starting in the first part with the spectra of systems of interacting particles, going on to the statistical mechanics of plasmas, and the amplification and attenuation of waves. In the second part, the subject of magneto-hydrodynamics is reviewed, and certain aspects of the stability of pinches are studied in considerable detail. The third part studies the relationship between microwaves and plasma physics, going into the question of shock waves, plasma vortices, power absorption, magnetron theory, and the use of microwaves to investigate the processes that take place in controlled fusion research. Finally, there is a chapter describing Plasma Physics Research at M.I.T.

This book, representing the work of recognized authorities on the various subjects treated, helps to provide a systematic exposition of many aspects of plasma theory, most of which have developed so recently that they are available only in scattered articles in the published literature. It also contains some new material that is not to be found elsewhere. On the whole, the presentation of the various articles seems to be satisfactory. As is perhaps inevitable in a collection of this kind, the over-all effect is somewhat disjointed. The book therefore does not serve as a useful introduction to plasma physics, of a kind that is needed, for example, by a student who is just beginning in the field. On the whole, the book is more likely to be useful to those who already have a good general grounding in the subject, and who wish to go into it more deeply. D. J. Bohm

Wave propagation in a turbulent medium. By V. I. Tatarski, translated by R. A. Silverman. (New York, Toronto, London: McGraw-Hill, 1961.) Pp. xiv + 285. Price 76*s*.

This monograph is a comprehensive account of the scattering and modification of electromagnetic and sound waves as they pass through a medium with random fluctuations of velocity and density caused by turbulent flow. The material is predominantly theoretical but there is included a considerable quantity of experimental measurements that test the theoretical predictions. The first of the four parts opens with a brief account of random functions and random fields and of their statistical description in terms of correlation functions, spectrum functions and structure functions, and uses these methods for the description of the random fields of velocity and of passive additives to a turbulent flow. The theory of local isotropy, which is particularly relevant to the subject of the monograph, is developed and the most important results for velocity and temperature fluctuations are obtained. The second part contains the theory of scattering of electromagnetic and sound waves by a finite volume of fluid containing time-independent, random fluctuations of velocity and temperature. In an appendix, R. H. Kraichnan points out an error in the approximate equation used in the calculations for sound waves. The third part deals with the modification of waves travelling through the random fluctuations in the atmosphere, first using the approximation of geometrical optics and then diffraction theory for the more common situations in which diffraction is not negligible. Intensity fluctuations of star images, of light and sound beams and phase fluctuations of sound beams are considered. The last section presents a considerable quantity of experimental measurements, both of the temperature fluctuations in the atmosphere and of the effects of these fluctuations on waves, comparing the observations with the preceding theory.

The outstanding feature is the comparable degree of attention paid to the two aspects of the problem, the nature of the turbulent fluctuations and the effect of them on the wave propagation, and, in spite of the number of equations written out, the relation of the theory to practical problems is emphasized continually. For this balanced approach and for the inclusion between two covers of a comprehensive account of the subject, it can be recommended to anyone interested in the propagation problem and it is also interesting to those interested in the turbulence problem. In particular, the comparison of experimental measurements with the theoretical predictions is a considerable addition to the evidence favouring the validity of the Kolmogoroff theory of local isotropy, evidence which is not as convincing as is sometimes believed. A. A. Townsend

Einführung in die Theorie der Abtastsysteme. By J. Tschauner. (Munich: Oldenbourg, 1960.) Pp. 185. Price DM 32.

This is a book on the theory of Sampled Data Systems, a subject of great engineering interest. It has found several able expositors in the English language, such as Martin L. Shooman in Mishkin-Braun *Adaptive Control Systems*. Nevertheless, were it not for the language one could expect Tschauner's book to become very popular because it is an exemplary clear and lucid exposition of the whole subject starting from first principles, with all the mathematical tools, and leading up to problems of considerable complication. It would be desirable for a British publisher to arrange for an English translation. D. Gabor

Vacuum physics

Vacuum technology transactions. Proceedings of the Sixth National Symposium. Edited by C. R. Meissner. (Oxford, London, New York, Paris: Pergamon Press, 1960.) Pp. xvi + 335. Price £6.

The annual conference of the American Vacuum Society which started in 1954 has now become a well established event. The contributions to the early conferences were of varying standard, but each year has seen the general quality of the papers rise until now a British vacuum physicist or engineer must regard with envy the advantages his American counterpart has for scientific interchange of information. Over 900 people attended the 1959 symposium to hear 58 papers. Your reviewer attended this year's conference where the papers had risen to some 70 and the attendance to over 1000, of which about 70 came from Europe. In

spite of the large attendance the conference proceeded smoothly and was admirably organized.

The transactions under review are divided into seven sections and from these one can see the new trends in vacuum engineering and physics. Thus apart from the continued attention given to vacuum pumps, instruments, thin films, etc., there are sections on gas analysis, ultra high vacuum systems, ionic pumping and electron bombardment sources of novel design for melting and evaporating metals. One can observe the application of vacuum physics to space research with the development of large high vacuum systems for component evaluation.

Some of the papers which were of particular interest to the reviewer are briefly mentioned here. Dayton has developed relations between pumping speed requirements and the degassing rates of materials in a serious attempt to place system calculation on a firmer basis than the rough estimation at present in use. Much work has been done recently in determining the degassing rates of materials used in vacuum engineering, but there is need to extend the investigations and unify the methods of presenting data. The paper by Klopfer and his co-workers on the nature and reactions of residual gases in sealed off vacuum systems should be of particular interest to those concerned with sealed electronic apparatus. Work is described at the National Bureau of Standards on rendering gas beams visible using electron scattering in a similar manner to the Schlieren optical method. This latter technique should be a valuable aid to anyone concerned with the emission of gas beams into vacuum. It would be useful (particularly for the reviewer!) if all of the papers in future transactions had abstracts.

A limited number of the contributions in the transactions have appeared in the journals and whilst it is often difficult when reporting at conferences not to re-use work which has been previously published, the reviewer considers dual publication of identical material is undesirable. However, the contributors are not entirely to blame for this development. Publication of the transactions has usually been a year after the conference and understandably a contributor does not like to see the results of his work laying dormant, whilst other people have had the advantages of its public disclosure. The reviewer knows that the American Vacuum Society are taking every precaution to ensure early publication of the current conference and we can only wish them through their editor, C. R. Meissner, every success in their valuable work.

L. HOLLAND

Molecular distillation. By G. BURROWS. (Oxford: Clarendon Press, 1960.) Pp. viii + 214. Price 35*s*.

Some yea rs ago the reviewer advanced the view that there was need for works on specific topics in high vacuum physics and engineering in place of the continuous stream of books which treated the subject generally. In recent years the situation has improved and *Molecular Distillation* is a welcome addition to a growing range of texts by specialists. Mr. Burrows has been concerned for some years with the distillation of materials in vacuum and was associated with Dr. C. R. Burch in his early investigations on the oil diffusion pump and low vapour pressure fluids.

The subject matter of the book covers the theory of distillation in vacuum, the design of molecular stills and degassing apparatus, and the uses of the process.

The initial discussion on the evaporation process is interesting and Mr. Burrows directs new attention to the problem of calculating evaporation and condensation rates for systems where the mean free path of vapour molecules in the emitted vapour is dependent on their direction of propagation. This section is undoubtedly worth-while reading for all those concerned with evaporation processes at low gas pressure and who have come to accept the simple interpretation of most physics texts.

As is usual practice in a work on distillation Raoult's law is used for finding the partial vapour pressures of the components of a mixture. However, here, as with other systems such as metal alloys, one finds that the law is limited in its application because the departure from the ideal system is rarely known factually in terms of the activity coefficients. The reviewer has been concerned with the evaporation of inorganic materials in vacuum and cannot speak dogmatically about molecular distillation but he believes that this book should be a stimulant to the establishment of more experimentally determined data relating to real systems.

The discussion on rate of evaporation and its relationship to vapour pressure on pages 11 and 12 could be improved. It is stated that the number of molecules escaping from a liquid in vacuum is the same as the number arriving at the liquid when it is in equilibrium with its vapour, but the discussion might have been amplified to explain that in the presence of surface contamination the number of molecules evaporating are equal to the number which condense on striking the liquid. Hickman has shown that non-volatile contaminants on static surfaces can have a great effect on evaporation rates. Thus the author might have introduced an evaporation (or condensation) coefficient in the formula relating vapour pressure to evaporation rate.

There are one or two trivial criticisms about the treatment of high vacuum apparatus. Thus the diagrams for trapped 'O' ring seals show rectangular grooves whereas in practice it is often preferable to use re-entrant grooves for holding the gaskets in position. In the discussion on diffusion pumps the performance is given of a range of commercial pumps with backing pressures of about 0·2 mm Hg. This is misleading since diffusion pumps have been available for many years with built-in ejector stages which give backing pressures of 0·5 mm Hg. Vacuum systems used for pumping volatiles invariably require a high backing pressure so that the rotary pump can be fully gas-ballasted. Similarly a reference to the self-purifying system used on some diffusion pumps for removing contaminants from the boiler would have been as useful as the treatment given to fractionating type pumps.

Most of the foregoing objections are not of fundamental importance because a student can always consult a general work for information on vacuum components but he will require Mr. Burrows' book for a good introduction to molecular distillation.

L. HOLLAND

Measurement of metal transfer in electrical contacts by the radioactive isotope method

by F. LLEWELLYN JONES, M.A., D.Phil., D.Sc., F.Inst.P., M. R. HOPKINS, M.Sc., Ph.D., F.Inst.P., and C. R. JONES, B.Sc., Ph.D., A.Inst.P., University of Wales, University College of Swansea

MS. received 21st March 1961

Abstract

The use of radioactive isotopes in the measurement of the transfer of metal from one electrode of an electrical contact to the other is described and details of the experimental procedure are given. The relation between matter transfer and circuit inductance at very low values of the inductance has been determined for platinum and palladium contacts operating at potential differences of the order of one volt. It is shown that for these metals there is no range of inductance down to 10^{-8} H over which the transfer is independent of inductance. The amount of transfer in relation to the volume of the molten metal bridge between the electrodes is considered, and the significance of the results in the light of theories of the phenomenon of transfer is discussed.

1. Introduction

ELECTRICAL contacts may be conveniently classified into three main groups. First, there are the heavy duty contacts of power engineering; these may carry very large currents and the potentials may be thousands of volts; secondly, the light duty contacts widely used in telecommunication systems carry currents of a few amperes at potentials of the order of a volt; in the third group are electrostatic contacts in which potentials and currents are extremely small. It is with contacts of the second group that this communication is concerned. Such contacts eventually fail to operate owing to the transfer of metal from one electrode to the other, leading to the appearance of a crater in the one and excess metal on the other, which interfere mechanically with the operation of the contact.

Metal transfer in electrical contacts has been the subject of much investigation for many years; an account and assessment of this work has been given in previous publications (Llewellyn Jones 1953, 1957) in which the significance of experimental data on transfer, in relation to the theory of the transfer mechanism, has been discussed. Quantitative data on transfer can provide important evidence concerning the role played by the microscopic molten metal bridge found at the separating electrodes of a contact. For example, it is well known that, in a symmetrically formed molten metal bridge established between monometallic contact surfaces, temperature asymmetry can be set up by the Thomson effect, the section of highest temperature being displaced from the position of geometrical symmetry. If the bridge breaks at the region of highest temperature, then effective transfer of matter must result, and the amount transferred can in principle be computed from the geometry of the bridge. In general, the temperature displacement in a bridge is small, because of the smallness of the Thomson coefficient, so that the volume of metal transferred should, on this view, be a small fraction of the total volume of the bridge. The measurement of the actual volume of metal transferred compared with the total volume of the contact bridge is therefore of considerable physical significance in the elucidation of the mechanism of transfer.

The measurement of metal transfer presents some difficulty owing to the very small amount of metal transferred in a single operation of the contact. Weighing the electrodes to determine the transfer requires something like 10^6 operations of the contact before any useful estimation can be made. The gravimetric method has also the disadvantage of determining only the net transfer, because it is insensitive to the details of the exchange of metal between the two electrodes. Methods involving optical estimation at high magnification are subject to the same disadvantages. Many measurements of transfer have been made without full realization of the important role played by circuit inductance and the significance of these measurements must therefore be assessed with some care.

It has been pointed out previously (Llewellyn Jones 1953) that the physical conditions, at the electrode surfaces, and in the gas between them, immediately after the microscopic molten metal bridge has exploded, are highly conducive to the establishment of an arc discharge provided sufficient voltage is available. It is for this reason that the value of the local inductance is of great physical significance, and the fact that inductances very much less than a microhenry can in ordinary circuits set up potential differences of about 10 v or so, which are adequate to produce an arc, is of considerable importance—an importance which was not fully appreciated in the earlier work on measurements of metal transfer. Consequently, in any investigation of the metal transfer between contact electrodes immediately after the breaking of the bridge it is essential that the local inductance should be controlled and measured down to values less than 10^{-7} H and preferably less than 10^{-8}H. The static circuit voltage may be low—too low to initiate or maintain an arc—but the actual transient contact voltage on breaking the contact can be much higher on account of the local self inductance.

The object of the present paper is to describe a quantitative study of the transfer of metal from one electrode of an electrical contact to the other under closely controlled conditions of circuit inductance, by a method which is a considerable improvement on optical and gravimetric methods. One electrode is subjected to neutron bombardment and the migration of a radioactive isotope from that electrode to the other is measured.

By this means it is possible to obtain the relation between

transfer and inductance over a range of inductance extending down to 10^{-8} H. Care has been taken to confine the investigation to that of phenomena occurring at the breaking of a contact, these being separated from the less pronounced effects which are possible at the making of the contact.

2. Experimental method and procedure

2.1. *Method*

The principle of the method is to use an electrical contact having one radioactive electrode, and to measure the radioactivity of the other after the contact has been made to perform a known number of breaking operations. One of the electrodes of the experimental contact was therefore sent to be irradiated with neutrons in a nuclear reactor at the Atomic Energy Research Establishment at Harwell. In this way about one in 10^9 of the atoms of the irradiated material was converted into a radioactive isotope. The subsequent disintegration of this isotope resulted in the emission of β and γ rays. The irradiated electrode was then replaced in the contact system and the amount of metal transferred to the inactive electrode in a number of operations determined. The amount of metal migrating in the opposite direction was determined by reversing the polarity of the electrodes. The sensitivity of the method is such as to reduce the number of operations required for a reliable determination by a factor of about 500 as compared with methods involving weighing or optical estimation.

2.2. *Apparatus*

The simple circuit arrangement illustrated in Fig. 1 was designed to ensure that the metal transfer actually measured was that due to the breaking and not to the making of the circuit. Some such arrangement is essential because of the tendency of the electrodes to touch several times during the complete 'making' operation, rebounding slightly between each closure.

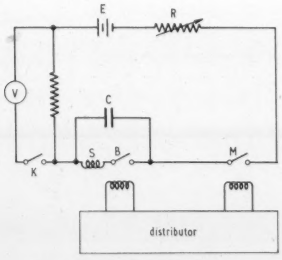


Fig. 1. The contact circuit.

B represents the experimental contact under investigation, M a separate relay contact, both operated by relay coils connected to a distributor unit having a rotating shaft with spring loaded contacts bearing on it. The distributor contacts are arranged so that coils are energized so as to perform the sequence: make B, make M, break B, break M, any desired number of times, the number of sequences being counted by a revolution counter. E is the low voltage battery supplying currents up to 35 A and C a 4 μF condenser connected across the contact. The inductance of the circuit was varied for the purpose of the experiments by plugging in coils S of inductance ranging from 6×10^{-8} H to 6×10^{-5} H in series with M. The coils were specially constructed of thick copper wire. Inductance was measured by means of a Marconi Q-meter having an oscillator of frequency range from 50 kc/s to 50 Mc/s. The residual self-inductance of the contact was 10^{-8} H. The electrodes of the contact M consisted of high purity metal disks $\frac{1}{8}$ in. in diameter and $\frac{1}{16}$ in. thick, butt soldered on to $\frac{1}{2}$ in. lengths of $\frac{3}{16}$ in. diameter copper rod. The speed of separation of the electrodes was between 0·5 cm/s and 1 cm/s.

To avoid danger from the small amount of radioactive vapour produced during the operation of the experimental contact, the contact was fitted into a steel chamber having an extractor fan and filter. The experimenter was protected at all times from radiation by adequate shielding.

2.3. *Detection and counting equipment*

The determination of transfer of radioactive material from one contact electrode to the other was made by means of a conventional Geiger–Müller counting system. This consisted of the Geiger–Müller counter itself, a stabilized power unit supplying the high voltage needed to operate the counter, an amplifier to amplify the pulses produced by the counter and a scaler which registered the pulses received. The choice of counting tube depended upon the radiation being measured, and the tube used in these experiments was of the type 2B7, manufactured by the General Electric Company, and having a thin aluminium window 7 mg/cm^2 in weight and 7 cm in diameter. The power unit, amplifier and scalar were made by Dynatron Radio Ltd. The unit was capable of supplying potentials up to 4 kv, maintaining a 0·5% stability for variation of input between 205 and 245 v.

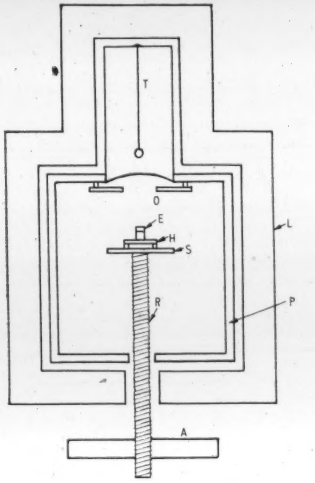
To avoid counting spurious pulses due to the release of secondary electrons from the cathode of the counter after each discharge, the scaling unit could be rendered inoperative for a predetermined period of between 1 μs and 100 ms after each operating pulse. For the G–M tubes which were used in these experiments this paralysis time was set at 200 μs. The maximum speed of counting was 500 counts per second.

The arrangement of the radioactive electrode under investigation in relation to the counting tube was such as to satisfy three requirements. The electrode had to be mounted in exactly the same position with respect to the tube each time a count was made. The background count had to be reduced to a minimum, particularly when dealing with weakly radioactive substances, and scattering within the apparatus had to be kept as low as possible. The scheme adopted is illustrated in Fig. 2. The cathode flange of the G–M tube was securely clamped in a circular aperture cut in the top side of a Perspex box. Perspex ($\frac{1}{8}$ in. thick) was used because its scattering factor is very low. An $\frac{1}{8}$ in. thick brass plate with a 1 in. orifice at its centre was mounted 0·5 cm beneath the tube window, the orifice being concentric with the anode wire of the tube. This defining orifice ensured that the rays entered the most sensitive part of the tube.

The electrode E was placed on a Perspex sample carrier S consisting of a platform whose vertical position could be adjusted by rotating the rod R which was threaded and screwed through the thick aluminium plate A. The position of the electrode was adjusted to a fixed height measured by a travelling microscope. The Perspex box P containing the arrangement was surrounded by a lead box with $\frac{1}{2}$ in. thick sides. This reduced the background count from about 120 to 50 counts per minute.

2.4. *Corrections to be applied to experimentally determined counts*

In measurements of radioactivity by means of counters, comparison of observed counts is meaningless unless the various factors affecting the magnitude of a count are taken into consideration. These are: (i) the paralysis time of the counting equipment, (ii) the time decay of the radioactive material, (iii) absorption in the air and in the window of the counter, (iv) the geometry of the counting system, (v) scattering from neighbouring objects into the tube, (vi) scattering from the non-radioactive part of the object under investigation, and (vii) absorption in the source itself.



T Geiger-Müller tube O Defining orifice E Electrode
H Electrode holder S Horizontal platform P Perspex lining
L Lead shielding R Threaded rod A Aluminium plate

Fig. 2. Geometrical arrangement of counting system.

In the method used in these experiments correction factors resulting from (iii), (iv), (v) and (vi) were the same for all deposits. Since no absolute determination of the radioactivity was required, these corrections were not applied. Corrections for (i), (ii) and (vii), however, are relevant. Correction (i) to account for paralysis time is simple. If the paralysis time is m seconds, and a count of n pulses per minute is observed, then the corrected count is $N = n/(1 - mn/60)$. This correction was applied to all observations.

To correct for (ii), the decay of activity of the radioactive material with time, the following procedure was adopted. Several contact electrodes of the same metal and of identical size were prepared. Calling them a, b, c, d, . . ., electrode a was sent to Harwell to be irradiated. This electrode was then used to deposit a layer of radioactive material on b, by using a and b as a pair in a contact for a suitable number of operations; b was then kept as a standard, its count being taken at regular intervals and recorded as a function of time. Electrodes c, d, . . ., were the ones used with a in the experimental investigation of transfer. By comparing the count of a deposit obtained on, say, c, from a in a transfer experiment, with the count of the standard b at that time it was possible to determine the count the deposit would have given at any other time. In this way all counts were corrected to a standard time so that the corresponding amounts of material were immediately comparable.

To correct an observed count for absorption in a specimen of thickness t, the observed count was multiplied by the factor $\mu t/(1 - e^{-\mu t})$, μ being the linear absorption coefficient of the material of the specimen (see, for example, Taylor 1951). In the experiments described here this factor was nearly unity, so that it was sufficient to estimate t crudely from the observed count and the observed diameter of the approximately circular deposit. Since μ depends on the density of the absorbing material, it was determined for aluminium by using successive layers of aluminium foil to determine the total thickness of foil required to reduce a given count to half its value. μ for the bulk electrode material was estimated by multiplying by the ratio of densities. It was then assumed that the density and therefore absorption coefficient of the radioactive deposit was the same as that of the bulk material. The magnitude of this correction for absorption in the specimen was from 2% to 5%.

2.5. *Determination of the volume of a radioactive deposit*

The volume of radioactive material deposited on a contact electrode in a transfer experiment was determined by comparing its count with the count of a small piece of comparator foil. This foil was of the same material as the electrodes and was irradiated in the nuclear reactor under the same conditions as the radioactive electrode, so that its specific activity at any time was identical with that of the electrode. The volume of the foil was accurately measured by means of a microscope and its count observed over a known period. In order to make the conditions identical with those for ordinary deposits, the count was taken with the foil placed on top of an electrode. This ensured that the back-scatter conditions were the same as those in the transfer measurements. After correcting for self-absorption in the comparator foil the count per unit volume was calculated. This was then used to estimate the volume of any radioactive deposit.

3. Results of transfer experiments with platinum and palladium

3.1. *Experimental Procedures*

3.1.1. *Isotopes*

The production of radioactive platinum involves the conversion of ^{196}Pt to ^{197}Pt, the other isotopes produced in the reaction being ^{193}Pt, ^{199}Pt and ^{199}Au. The isotope ^{197}Pt has a half-life of 18 hours. Experiments depending upon the activity of ^{197}Pt would be difficult because of its short half-life. The isotope ^{199}Pt, however, decays rapidly, with a half-life of 1 hour, to form ^{199}Au of high activity and a half-life of 3·4 days. ^{193}Pt has a half-life of 4·33 days. In a set of experiments requiring about ten days to complete, the isotopes of interest were ^{199}Au and ^{193}Pt.

Palladium subjected to neutron bombardment forms two radioactive isotopes, namely ^{109}Pd and ^{111}Ag. The nuclear reactions giving rise to these isotopes are

$$\text{(i)}\quad {}^{108}\text{Pd} + \text{n} = {}^{109}\text{Pd} + \gamma$$

and

$$\text{(ii)}\quad {}^{110}\text{Pd} + \text{n} = {}^{111}\text{Pd} + \gamma$$
$$\qquad\qquad {}^{111}\text{Pd} = {}^{111}\text{Ag} + \beta.$$

The isotope ^{111}Pd which is formed in the production of ^{111}Ag has a half-life of 22 minutes and does not survive for long after the palladium is removed from the source of neutrons. ^{109}Pd has a half-life of 13·6 hours and ^{111}Ag a half-life of 7·5 days. The experiments using the activated palladium were begun two days after its removal from the atomic pile. After this time there was little of the ^{109}Pd left, so that the activity of the palladium during the following fortnight in the course of which the experiments were being performed, was due to ^{111}Ag. The proportion of palladium transformed to silver is exceedingly small (1 in 10^9) so that it does not affect the character of the palladium.

3.1.2. *Experimental results*

The results of measurement of transfer of platinum as a function of the self-inductance L, at a current of 15 A and an open circuit voltage of 4 V are illustrated in the graphs of Fig. 3. Preliminary results have already been briefly reported (Llewellyn Jones 1957). The first graph shows the migration

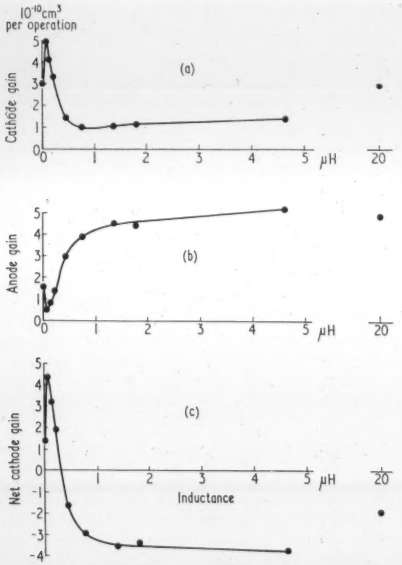


Fig. 3. Transfer in a platinum contact (current = 15 A).

of anode metal to the cathode in a contact consisting of a radioactive anode and a non-radioactive cathode. The graph thus shows the variation of cathode gain with inductance. The second graph shows the migration of cathode metal to the anode, the cathode in this case being radioactive and the anode not. The third graph is obtained by taking the difference between corresponding ordinates in the first and second, and thus represents the net transfer to the cathode over a range of values of the inductance. Each of the curves (*a*) and (*b*) is almost a mirror image of the other. By adding corresponding ordinates a measure of the amount of metal which is not on its original electrode is obtained. From an inductance of 1×10^{-8} H to 2×10^{-6} H this appeared to be constant to within $\pm 10\%$, indicating that the total amount of metal involved in the transfer process remained constant. The volume of the cathode deposit formed by 1000 operations was estimated to be 5×10^{-7} cm^3 for $L = 6 \times 10^{-8}$ H and a current of 15 A, giving a value of 5×10^{-10} cm^3 per operation.

The corresponding results for palladium are shown in Fig. 4.

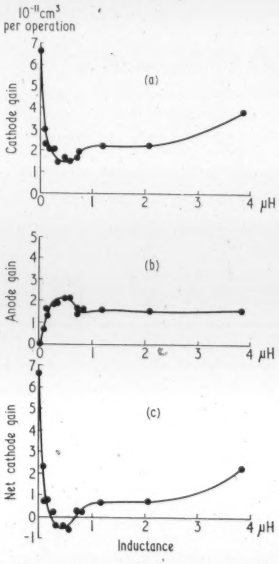


Fig. 4. Variation of transfer with inductance for palladium (current = 15 A).

3.2. *The relation between volume of metal removed from both electrodes and volume of the molten metal bridge*

The effect of a single operation of a platinum contact is to produce on each electrode a smooth circular crater; surrounding each crater there is a ring of projecting metal. The appearance of the craters is such as to suggest that the molten metal is blown out of them and deposited on both electrodes. Photographs illustrating this effect have been published previously (Llewellyn Jones 1957, Llewellyn Jones and Jones 1957). Let it be assumed at this stage that the total picture is consistent with the view that the surfaces of the craters are the isothermal surfaces separating the molten metal of the bridge from the solid electrodes. The shape of each crater is approximately that of part of a hemispherical cup. The total volume of the molten metal bridge is then equal to that of the two craters from which the metal between the two isothermal surfaces must have come.

The sum of the volumes of the anode and cathode craters after a single contact operation at a current of 15 A was estimated by the optical measurement of the erosion patterns at high magnification to be 8×10^{-10} cm^3. It is important to compare this estimate of the volume of the molten metal bridge with the sum of the volume of metal removed from anode to cathode and that removed from cathode to anode

as determined by the radioactive isotope technique, at the same current and over the range of inductance for which it is constant (i.e. 1×10^{-8} H to 2×10^{-6} H). The volume moved from anode to cathode in one operation at an inductance of 6×10^{-8} H was 5×10^{-10} cm^3 per operation and the volume moved from cathode to anode $0{\cdot}5 \times 10^{-10}$ cm^3, the total being $5{\cdot}5 \times 10^{-10}$ cm^3; the net transfer to the cathode was thus $4{\cdot}5 \times 10^{-10}$ cm^3. In platinum under these conditions, therefore, most (in this particular case, three-quarters) of the metal of the bridge was moved in either direction across the gap between the electrodes. The net transfer to the cathode was about half the volume of the bridge.

Theories of metal transfer based on temperature asymmetry in the molten metal bridge between the contact electrodes suggest that it should be possible to establish conditions at very low inductance which would leave temperature distribution in the bridge to be the determining factor in the transfer process. This means that the volume of metal transferred would be independent of the circuit inductance provided this inductance were low enough. Warham (1953), for example, on the basis of some observations of transfer by optical means, concluded that the transfer was independent of inductance up to an inductance of 7×10^{-6} H in the case of platinum. The results of the present investigation using the radio-tracer technique show that there is no range of values of the inductance over which the transfer is independent of the inductance, even down to an inductance of 10^{-8} H. This suggests that explanations of transfer depending only on temperature asymmetry in the bridge are inadequate.

The results also show that there is a range of inductance (1×10^{-8} H to 2×10^{-6} H for platinum, under the experimental conditions described above) over which the total amount of metal involved in the transfer process is constant. This amount of metal approaches that of the whole of the molten metal bridge. At higher values of the inductance the amount of metal is increased owing to the occurrence of a sustained arc discharge.

4. Conclusions

The main conclusions which have been reached as a result of the measurements of transfer apply equally well to platinum and palladium. The radioactive isotope technique demonstrates the practically simultaneous transfer of matter to each electrode independently. This exchange could not have been observed either by optical or gravimetric methods. The circuit inductance plays an important part in the direction of transfer right down to the lowest values investigated. While in the case of platinum the total amount of metal transferred from its original electrode was comparable with the total volume of the molten metal bridge between the electrodes, in the case of palladium, on the other hand, it was only a small fraction. The significance of these results must now be discussed in relation to the mechanism of metal transfer.

A process which depends on a small displacement of the region of maximum temperature as a result of thermoelectric effects in the bridge would cause the transfer of only a small proportion of the material of the bridge. The transfer of most of the bridge would be difficult to understand on this view, and some mechanism other than the thermoelectric shift must be considered. In the case of palladium the transferred material is only a small fraction of the volume of the bridge and this does not preclude the thermoelectric explanation. But the result, which applies to both metals, that the transfer is never independent of the inductance, shows conclusively that the thermoelectric mechanism is not the determining factor in the transfer process. In platinum, for example, the observations are consistent with the view that the molten metal bridge is completely dispersed as liquid droplets which are distributed between the electrodes. The droplets, at high temperatures, would acquire a positive charge due to the emission of electrons and, in the presence of the contact field, would be drawn towards the cathode, except for the proportion which might be blown on to the anode by the explosive breaking of the bridge. The appearance of a ring of raised metal around the erosion crater (Llewellyn Jones and Jones 1957) suggests that the exploded matter moves outwards from the axis of the bridge as well as from electrode to electrode. A tentative theory based on the movement of droplets was in fact proposed by Bühl (1926).

When, however, one takes into account the fact that the net transfer is dependent on the circuit inductance, the force on the particles travelling from one electrode to the other must depend on the transient field and therefore the potential difference across the contact at break. It is difficult to understand the movement of drops of any size under the conditions of rapid oscillation that ensue. It is easier to reconcile the phenomena with the view that the exploding bridge gives rise to a plasma containing metal vapour at high pressure and that movement of matter takes the form of ionic drift in the high electric field between the electrodes of the breaking contact.

Acknowledgments

We wish to thank the Royal Society for a grant to enable us to obtain precious metals of very high purity for this work. We also wish to thank the Department of Scientific and Industrial Research for a maintenance grant to one of us (C. R. J. 1954–57), and R. L. Griffith for his help in some of the later measurements.

References

BÜHL, A., 1926, *Ann. Phys.*, **80**, 132.

LLEWELLYN JONES, F., 1953, *Fundamental Processes of Electrical Contact Phenomena* (London: H.M.S.O.).

—— 1957, *The Physics of Electrical Contacts* (Oxford: Clarendon Press).

LLEWELLYN JONES, F., and JONES, R. H., 1957, *Z. Phys.*, **147**, 543.

TAYLOR, D., 1951, *The Measurement of Radio Isotopes* (London: Methuen).

WARHAM, J., 1953, *J. Instn Elect. Engrs*, Pt I, **100**, 163.

Experimental determination of the wall effect for spheres falling axially in cylindrical vessels

by V. FIDLERIS,* B.Sc., Ph.D., A.Inst.P., and R. L. WHITMORE, B.Sc., D.Sc., F.Inst.P., Department of Mining and Fuels, University of Nottingham

MS. received 3rd *March* 1961, *in revised form* 13th *April* 1961

Abstract

Experimental data of the drag exerted by the walls of a cylindrical vessel on a sphere falling axially down it through a liquid are given for Reynolds numbers, based on the diameter of the sphere, between 0·05 *and* 20000. *Existing wall-correction formulae are examined in the light of the new data, the conclusion being that the Francis* (1933) *and Munroe* (1888) *equations are the most reliable in the laminar- and turbulent-flow regions respectively. Graphs show the correction to be applied in the intermediate-flow region.*

1. Introduction

THE object of the work was to make reliable measurements of the terminal velocities of spheres falling axially in cylindrical vessels, under conditions extending from purely-laminar to fully-turbulent flow, and to determine the correction to be applied for the retarding effect of the walls on the falling spheres. Available literature relating to the wall effect is both numerous and of high quality (Newton 1687, Munroe 1888, Ladenburg 1907, Sheppard 1917, Faxen 1921, Lunnon 1928, Schmiedel 1928, Barr 1931, Francis 1933, Fulmer and Williams 1936, Gurel 1951, Mott 1951), but most of it is limited to the laminar-flow region. No wall corrections appear to exist for the intermediate-flow region (Reynolds numbers from 1 to 500) and the few published results in the turbulent region (Newton 1687, Munroe 1888, Lunnon 1928, Mott 1951), are widely scattered. In order to fill this gap, experiments were made with 60 spheres (covering a range of diameters and densities), four cylinders of different diameters, and fifteen Newtonian liquids of various viscosities and densities. The range of flow covered Reynolds numbers from 0·054 to 20000 and the results were used to assess the validity of known wall-correction formulae.

* Now at Atomic Energy of Canada Ltd., Chalk River, Ontario.

2. Experimental

An indirect method of measuring the falling speed of the test spheres was employed and a detailed description of the apparatus and its accuracy can be found elsewhere (Fidleris and Whitmore 1959). Essentially it consisted of a copper reservoir, in which the liquid was brought to the required temperature, and a jacketed vertical glass cylinder, in which the time of fall of the test spheres over a known distance was measured. The temperature of the water in the jackets of the copper reservoir and the experimental tube was carefully controlled by pumping it through coils immersed in a constant-temperature bath. When equilibrium was reached the temperature of the liquid and the jacketing water could be maintained to within ±0·05 deg C of the required temperature, which for most measurements was chosen to be 20·0° C.

The detection of the test spheres was based on the measurement of the change of impedance of coils, wound coaxially with the experimental column and spaced (30 ± 0·01 cm) apart, when a metallic test sphere passed through them. The time of fall of the test spheres was obtained from a photographic record of the detection signals. By comparison against a standard signal, the fall time could be determined to within ±0·3%. The top detection coils were situated 40 cm or more below the surface of the liquid to ensure that a test particle reached its terminal velocity before entering the detection system, which was sensitive enough to record the passage of a 0·05 cm diameter steel sphere or a 0·10 cm diameter sphere made of non-ferromagnetic material. The sphere-release mechanisms were designed to ensure that the test particles fell axially down the cylinders, which were precision-bore glass tubes 100 cm in length and 1·50 cm, 2·00 cm, 2·50 cm and 3·00 cm in internal diameter. The accuracy and constancy of tube diameters was in all cases better than ±0·1%.

The test spheres were made of metal or metal-coated plastic and, with the exception of four cases, were spherical to within a diameter variation of less than ±0·15%. Their details, including the reliability of the density determinations, are summarized in Table 1.

Table 1. *Test spheres*

No.	*Material*	*Density (g/ml.)*	*Diameter of smallest sphere (cm)*	*Diameter of largest sphere (cm)*
1	Iron-coated plastic	1·275 ± 0·002	0·495	0·554
2	Magnesium	1·800 ± 0·002	0·256	1·018
3	Magnesium–aluminium alloy	2·570 ± 0·005	0·256	0·763
4	Duralumin	2·805 ± 0·005	0·256	1·018
5	Steel	7·72 ± 0·02	0·0793	2·540
6	Lead	11·25 ± 0·03	0·251	0·374
7	Tantalum	16·60 ± 0·04	0·191	0·509
8	Gold	19·36 ± 0·02	0·100	1·000

Details of the liquids which were used in the experiments are given in Table 2. The density of each liquid was determined with a pycnometer to within an accuracy of 0·05%

Table 2. *Test liquids*

No.	Liquid	Density at 20° C (g/ml.) minimum	maximum	Viscosity at 20° C (cP) minimum	maximum
1	Water	0·998		1·005	
2	Olive oil	0·912		78·3	
3	Aqueous solution of glycerol	1·219	1·274	116	225
4	Aqueous solution of glycerol and lead nitrate	1·186	1·187	1·23	27·2
5	Aqueous solution of glycerol and *n*-propyl alcohol	1·055	1·186	40·2	86·4

and checked before each experiment with a standardized hydrometer. Their viscosities were originally determined with standard capillary instruments (as described in B.S. No. 188: 1937) and checked before each experiment with a Ferranti rotating-cylinder viscometer. The viscosity was considered to be the least reliable of all the measurements, possessing an error of ±1%.

3. Results

In all, some 3000 velocity determinations were made and the results can be found in detail elsewhere (Fidleris 1958). A typical example of the variation of the terminal velocity with the diameter of the test sphere and the cylinder is shown in Fig. 1, each point representing the average value of three

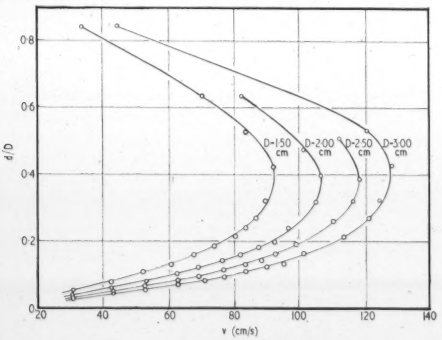


Fig. 1. Terminal velocity of steel spheres falling through water in tubes of various diameters.

or more velocity determinations. In order to summarize all experimental results, use was made of the logarithmic plot of the two non-dimensional quantities:

$$\text{Reynolds number,} \qquad (Re) = d\rho v/\eta \tag{1}$$

and

$$\text{resistance coefficient} \quad \psi = \frac{4g(s-\rho)d}{3\rho v^2} \tag{2}$$

where s = density of test sphere, ρ = density of liquid, g = acceleration due to gravity, d = sphere diameter, v = terminal velocity, η = coefficient of viscosity.

Fig. 2 contains the experimental results for a number of d/D ratios, where D is the diameter of the cylinder. They a smooth curve which broadly follows the ($\log \psi$, $\log (Re)$) curve for spheres falling through infinite fluid as determined by other workers (Wieselsberger 1922, Liebster 1927, Lunnon

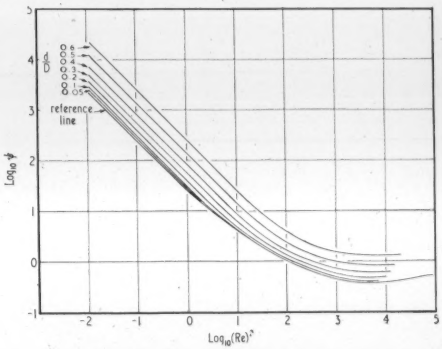


Fig. 2. Resistance coefficient–Reynolds number curves for spheres falling in cylindrical containers.

1928, Goldstein 1938, Möller 1938, Davies 1945). The scatter of the experimental points is such that 95% are within ±3% of the mean line.

4. Discussion of results

From Fig. 2 it is clear that the retarding effect of the wall on a falling sphere diminishes on transition from laminar to turbulent flow and decreases with increasing Reynolds number to an approximately constant value. Any correction to the falling speed of a sphere to allow for wall interference must, therefore, be a function of Reynolds number and in this paper a graphical method of arriving at the correction is suggested.

The first step is to relate the velocities of a sphere, falling through a fluid in cylinders of different diameters, to its terminal velocity in an infinite extent of the fluid. The Reynolds number may then be expressed as $(Re) = c_1 v$ and the resistance coefficient as $\psi = c_2/v^2$, where c_1 and c_2 are constants and depend on the properties of the sphere and the liquid used.

If the results are plotted logarithmically as in Fig. 2, an alteration in velocity from v_1 to v_2 due to a change in the diameter of the containing vessel is represented by a movement of the original point along a line of slope −2 for a distance a, equal to $\sqrt{5}\log_{10}(v_1/v_2)$ in log-scale units. If the

first point corresponds to a sphere falling in infinite medium, its velocity is v_∞. In a vessel of diameter D, the velocity is reduced by wall interference to a value v and $v_\infty/v = 10^{a l \sqrt{5}}$.

Conversely, if the test is made in a vessel of diameter D, it is only necessary to draw a line of slope -2 through the corresponding point on the (log ψ, log (Re)) curve to cut the curve for infinite medium (subsequently referred to as the reference line). The point of intersection represents the fall of the sphere in infinite fluid.

This procedure was applied to the curves plotted in Fig. 2. Because the present experimental investigations could not be extended to very large vessels, the results obtained by Wieselsberger (1922), Liebster (1927), Lunnon (1928), Möller (1938) and Davies (1945) were used to plot the reference line in Fig. 2. The points used in its construction are given in Table 3 and correspond to those quoted by Goldstein (1938). Fig. 3 shows the factors v_∞/v by which readings taken in a cylindrical vessel must be multiplied to give the terminal velocity in infinite medium. In Fig. 4 the corresponding v/v_∞ factors are shown for converting a sphere velocity, as calculated in infinite medium, to its falling velocity in a cylindrical vessel of given d/D ratio. The two corrections coincide in the laminar region and nearly so under full turbulence, but they differ significantly in the intermediate-flow region. This is because the Reynolds number is calculated from the uncorrected velocity and the slope of the log ψ–log (Re) curve alters appreciably in the intermediate-flow region.

Table 3. *Co-ordinates of the 'reference line' for spheres* ((log ψ, log (Re)) curve for infinite fluid)

(Re)	ψ	(Re)	ψ
0·01	2400	600	0·525
0·1	244	1000	0·455
0·4	63·0	2000	0·394
1·0	26·2	4000	0·380
4·0	8·20	7000	0·382
10	4·27	10 000	0·393
30	2·03	20 000	0·430
60	1·36	40 000	0·472
100	1·05	60 000	0·510
300	0·660		

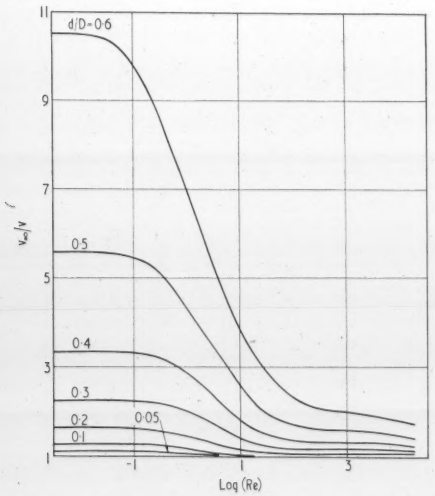


Fig. 3. Correction of fall velocity in a cylindrical vessel to that in infinite fluid.

5. Comparison with existing wall-correction formulae

The retarding effect of the walls of a cylindrical vessel on a sphere falling axially down it can be expressed in the form of a multiplying factor for correcting the falling velocity in an infinite medium v_∞ to the velocity v in a vessel possessing a particular d/D ratio.

An inspection of Fig. 2 suggests that the logarithmic (ψ, (Re)) curves are straight and parallel at Reynolds numbers less than unity, and very nearly so between Reynolds numbers of 1000 and 10 000. Thus under fully laminar or fully turbulent conditions, wall-effect equations which involve only d/D and not Reynolds number should be possible.

The former region has been extensively investigated by other workers (Ladenburg 1907, Sheppard 1917, Faxen 1921, Schmiedel 1928, Barr 1931, Francis 1933, Fulmer and Williams 1936, Gurel 1951) and a number of theoretical and empirical wall-effect equations have been proposed. Of them the three most widely used are those by Ladenburg (1907)

$$v/v_\infty = 1/\{1 + 2{\cdot}1(d/D)\} \tag{3}$$

which in turn is the first approximation of a theoretical equation derived by Faxen (1921), which for low Reynolds numbers may be expressed as

$$v/v_\infty = 1\text{–}2{\cdot}104(d/D) + 2{\cdot}09(d/D)^3 - 0{\cdot}95(d/D)^5. \tag{4}$$

Both of them have been shown (Barr 1931) to be limited in their range of application to small values of d/D, i.e. below 0·2 in the case of Eqn (4) and below 0·1 in the case of Eqn (3).

The empirical equation obtained by Francis (1933)

$$\frac{v}{v_\infty} = \left\{\frac{1 - (d/D)}{1 - 0{\cdot}475(d/D)}\right\}^4 \tag{5}$$

was found (Gurel 1951) to be satisfactory for d/D ratios of less than 0·9.

In Table 4, Eqns (3), (4) and (5) are compared with the experimental results of Fig. 4, at a Reynolds number of 0·1.

Table 4. *Comparison of wall-correction equations in the laminar region of flow* (v/v_∞)

d/D	*Experimental from Fig. 4*	*Ladenburg from Eqn* (3)	*Faxen from Eqn* (4)	*Francis from Eqn* (5)
0·05	0·894	0·905	0·895	0·898
0·10	0·788	0·826	0·792	0·797
0·20	0·610	0·704	0·596	0·611
0·30	0·446	0·613	0·423	0·444
0·40	0·300	0·543	0·283	0·301
0·50	0·180	0·488	0·179	0·185
0·60	0·094	0·442	0·115	0·098

It will be seen that the agreement is good in the case of the Francis equation and rather poorer using the Faxen equation.

In the turbulent region, at least five wall-effect equations have been proposed, but there is little published experimental evidence against which to test them. Newton (see

Table 5. *Experimental values of v/v_∞ in the turbulent region of flow*

	$(Re) = 100$		$(Re) = 1000$		$(Re) = 3000$		$(Re) = 10000$	
d/D	*from Fig. 3*	*from Fig. 4*	*from Fig. 3*	*from Fig. 4*	*from Fig. 3*	*from Fig. 4*	*from Fig. 3*	*from Fig. 4*
0·1	~1	~1	~1	~1	~1	~1	~1	~1
0·2	0·956	0·948	0·960	0·954	0·962	0·958	0·962	0·962
0·3	0·862	0·856	0·868	0·866	0·876	0·874	0·892	0·892
0·4	0·740	0·720	0·766	0·760	0·782	0·776	0·816	0·800
0·5	0·590	0·560	0·640	0·632	0·654	0·648	0·700	0·676
0·6	0·430	0·390	0·506	0·496	0·530	0·518	0·562	0·546

Table 6. *Comparison of wall-correction equations in the turbulent flow region* (v/v_∞)

d/D	*Experimental at* $(Re) = 3000$ *averages from Table 5*	*Newton from Eqn* (6)	*Munroe from Eqn* (7)	*Lunnon from Eqn* (8)	*Mott from Eqn* (9)	*Mott from Eqn* (10)
0·1	—	0·987	0·968	0·977	0·982	—
0·2	0·960	0·950	0·911	0·954	0·933	—
0·3	0·875	0·889	0·836	0·931	0·860	—
0·4	0·779	0·806	0·747	0·908	0·776	—
0·5	0·651	0·702	0·646	—	0·690	0·500
0·6	0·524	0·543	0·535	—	0·607	0·325

Barr 1931) made some calculations of the wall-effects of large spheres and obtained the correction term

$$\frac{v}{v_\infty} = \frac{(D^2 - d^2)\sqrt{\{D^2 - (d^2/2)\}}}{D^3}. \tag{6}$$

Munroe (1888) suggested the approximate formula

$$v/v_\infty = 1 - (d/D)^{3/2} \tag{7}$$

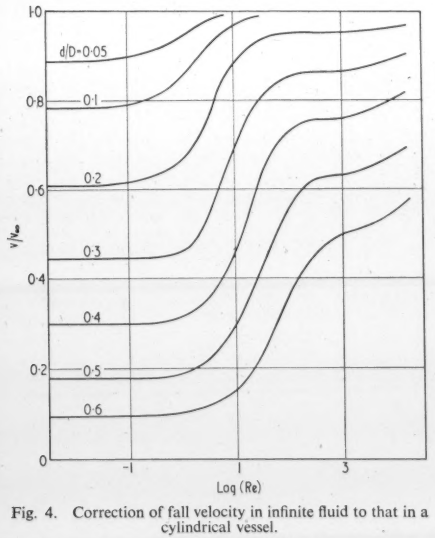


Fig. 4. Correction of fall velocity in infinite fluid to that in a cylindrical vessel.

which he based on measurements of the fall of lead shot in water. He found the wall-effect to be negligible for values of d/D less than 0·1. Lunnon (1928) obtained a series of results with steel spheres falling through air and water in four different tubes, and suggested a relationship

$$v/v_\infty = 1 - 0{\cdot}23(d/D) \tag{8}$$

which he assumed to hold up to a d/D ratio of 0·3. He did not attempt to derive a correction for higher values of d/D. Mott (1951) used Lunnon's results to give two more formulae. For values of d/D from 0·2 to 0·5 he obtained the equation

$$\frac{v}{v_\infty} = \frac{1}{1 + A(d/D)^2} \tag{9}$$

where A is a constant varying from 1·8 to 3·2, and for values of d/D from 0·5 to 0·7 the equation

$$\frac{v}{v_\infty} = \frac{1}{1 + (2d/D)^4}. \tag{10}$$

Mott considered the wall-effects to be negligible below a d/D ratio of 0·15.

In Table 5 the experimental values of v/v_∞, as obtained from Figs 3 and 4, are given for a range of Reynolds numbers and in Table 6 Eqns (6) to (10) inclusive are compared with the averaged results taken from Figs 3 and 4 at a Reynolds number of 3000. There is no close correspondence between any of the equations, although nearly all of them come in places to within 2% of the experimental results. Thus the agreement with Newton's equation (Eqn 6) is within 2% up to a d/D ratio of 0·3, but worsens considerably at higher ratios. The degree of correspondence of any given equation varies to some extent with Reynolds number, although not always in the same way. Munroe's correction (Eqn 7) shows best agreement with experimental results between Reynolds numbers of 1000 and 3000, whereas Newton's (Eqn 6) is most reliable at Reynolds numbers near 10000. The least satisfactory formula is that of Lunnon (Eqn 8).

6. Conclusions

The experimental results show that it is difficult to derive a single relationship to account for the change in the interference effect of the vessel walls on a falling sphere, which occurs with increasing Reynolds number. A correction can, however, be made graphically without difficulty.

It can be concluded that the retarding effect of the wall decreases with increasing Reynolds number, so that for ratios of d/D of 0·05 and 0·10, the wall correction becomes less than 1% for Reynolds numbers exceeding 5 and 30 respectively.

In the laminar-flow region (that is for Reynolds numbers not exceeding approximately 0·2) the Francis equation is the most satisfactory and in the range of d/D ratios from 0 to 0·4 the error does not exceed $\pm 0{\cdot}5\%$. It would therefore appear that in determinations of viscosity from observations of the velocity of a sphere falling axially down a cylinder containing the fluid, corrections for the retarding effect of the wall should be found from the Francis equation.

In the turbulent region the usefulness of available equations is limited, but Munroe's is most satisfactory at Reynolds numbers between 1000 and 3000, where an error of $\pm 2\frac{1}{2}\%$ may be expected for d/D ratios of less than 0·6. Newton's equation increases in reliability when the highest Reynolds numbers reached in the experiments, which were about 10000, are approached. In this region the error is only $\pm 1\%$.

In the intermediate region Fig. 3 can be used to transpose velocities measured in a cylindrical vessel to the corresponding velocity of fall of a sphere in infinite fluid, the Reynolds number being calculated from the uncorrected fall velocity in each case.

Acknowledgments

One of the authors (V. F.) wishes to thank the National Coal Board for financial support in carrying out the work.

References

BARR, G., 1931, *A Monograph of Viscometry* (Oxford: University Press).

DAVIES, C., 1945, *Proc. Phys. Soc.*, **57**, 259.

FAXEN, H., 1921, Dissertation (University of Uppsala).

FIDLERIS, V., 1958, Ph.D. Thesis (University of Nottingham).

FIDLERIS, V., and WHITMORE, R. L., 1959, *J. Sci. Instrum.*, **36**, 35.

FRANCIS, A. W., 1933, *Physics*, **4**, 403.

FULMER, E. I., and WILLIAMS, J. C., 1936, *J. Phys. Chem.*, **40**, 143.

GOLDSTEIN, S., 1938, *Modern Developments in Fluid Dynamics* (Oxford: Clarendon Press), p. 493.

GUREL, S., 1951, Ph.D. Thesis (University of Birmingham).

LADENBURG, R., 1907, *Ann. Phys., Paris*, **22**, 287; **23**, 447.

LIEBSTER, H., 1927, *Ann. Phys., Lpz.*, **82**, 541.

LUNNON, R. G., 1928, *Proc. Roy. Soc. A*, **118**, 680.

MÖLLER, W., 1938, *Phys. Z.*, **39**, 57.

MOTT, R. A., 1951, *Some Aspects of Fluid Flow* (London: E. Arnold and Co.), p. 242.

MUNROE, H. S., 1888, *Trans Amer. Inst. Min. Engrs*, **17**, 637.

NEWTON, I., 1687, *Principia*, Lib. II, Prop, viii, Cor. 2. Quoted by Barr 1931.

SCHMIEDEL, J., 1928, *Phys. Z.*, **29**, 593.

SHEPPARD, E., 1917, *Industr. Engng Chem.*, **9**, 523.

WIESELSBERGER, C., 1922, *Phys. Z.*, **23**, 219.

Epitaxy and twinning in foils of some noble metals condensed upon lithium fluoride and mica

by M. J. HALL and M. W. THOMPSON, B.Sc., Metallurgy Division, Atomic Energy Research Establishment, Harwell, Didcot, Berkshire

MS. received 9*th January* 1961, *in revised form* 30*th May* 1961

Abstract

Foils up to $2{\cdot}5 \times 10^{-3}$ *cm in thickness, of Cu, Ag, Pd and Au, have been evaporated on to LiF and mica crystals, epitaxial growth was observed within fixed temperature ranges for each substrate–metal combination. X-ray diffraction of the foils produced sharp patterns of spots which indicated a principal orientation of crystallites in direct alignment with the substrate crystal, together with twins whose presence was confirmed by metallography. The observations are discussed in terms of a nucleation and growth model.*

1. Introduction

It has often been shown that when a metal vapour condenses on a crystalline substrate within a limited temperature range, a film of metal forms with crystallite orientations closely related to the substrate crystal (for a review see Pashley 1956). This phenomenon is known as epitaxy. In most publications the film thickness has been in the range 10 to 1000 Å. A recent series of experiments (Thompson and Nelson 1960), in which sputtering from noble metals was studied, required single crystals in the form of thin foils about 10^{-3} cm thick. It was found possible to use polycrystalline foils containing a small number of well-defined crystallite orientations, and the account which follows describes and discusses their preparation by epitaxial growth on mica and lithium fluoride crystals.

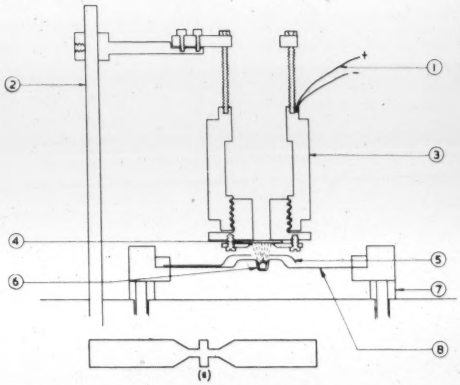


Fig. 1. Vertical section through the evaporation apparatus and furnace-block.

(*a*) Detail showing the construction of the filament used to heat the crucible.

1, Iron-constantan thermocouple; 2, rotating shaft; 3, copper furnace, slotted vertically, for heater windings; 4, substrate; 5, radiation and filament contamination shield; 6, alumina crucible and molten metal charge; 7, insulators 8, tantalum filament.

2. Apparatus

A copper furnace-block wound with nichrome tape (Fig. 1) was used to heat the substrate crystal in a vacuum into the temperature range where epitaxy occurs. The temperature was measured by an iron-constantan thermocouple. The substrate crystal was clamped by a washer to the underside of the furnace-block which was supported on a shaft passing through a Wilson seal in the wall of the vacuum vessel. This enabled the substrate to be moved over the vapour source whilst under vacuum. Some 200 mg of the metal to be evaporated was placed in an alumina crucible of 28 mm^3 capacity. The crucible was heated by a tantalum filament passing some 30 A, whose construction is shown in Fig. 1(*a*). A hole through the centre of the furnace-block enabled the substrate to be optically aligned with the crucible at the start of evaporation. A copper shield over the filament prevented contamination of the substrate by filament material and reduced the amount of heat radiated to the furnace-block.

3. Method of growth

The required substrate was clamped to the furnace-block, dust particles removed from its surface by a compressed air blast, and the system evacuated to below 10^{-4} mm Hg. The temperature was raised to a value depending on the substrate–metal combination, with the substrate not directly over the crucible. The metal charge in the crucible was then melted and allowed to outgas for a few seconds. The shaft through the Wilson seal was rotated to bring the substrate directly over the crucible and the evaporation rate adjusted so that an opaque film formed in a few seconds. When the metal charge had evaporated, after about 10 minutes, the filament and furnace-block heaters were switched off and when the temperature had fallen to 400° C a pressure of 1 mm Hg of helium was admitted to accelerate cooling.

After cooling, the foil was removed from the substrate by soaking in distilled water. In difficult cases nitric acid (or aqua regia for gold) was used instead of water. The foils produced in this apparatus were circular with an area of

$0{\cdot}6\ cm^2$ and thicknesses between $2{\cdot}5 \times 10^{-4}$ and $2{\cdot}5 \times 10^{-3}$ cm.

4. Substrate preparation

Three types of substrate were used: the {0001} cleavage plane of hexagonal muscovite mica, the {100} cleavage plane and {110}, {111} and {112} planes of cubic lithium fluoride. Cleavage of both mica and lithium fluoride was carried out with a safety razor blade. In lithium fluoride the last three planes were prepared by grinding and polishing in the following way.

A cleaved cube of lithium fluoride of 1 cm side was held on a grinding wheel, using 120 mesh carborundum grit with water as a lubricant, at such an angle as to produce approximately the required plane. When roughly half the cube was gone the exact plane was obtained with a 'Handimet' grinder using a combination square to set the angles relative to the cube faces. A fine polish was given to the surface using coarse and then fine alumina, finishing with 1 micron diamond dust. The crystal was cemented by this polished face to a metal block using Canada balsam. It could then be easily held whilst the other side was ground and polished parallel to the first, a final thickness of 1 to 2 mm being aimed at. The crystal was removed from the block by heating, cleaned with toluene, etched in 40% hydrofluoric acid and finally given a chemical polish by agitation in a bath containing 2 vol. % ammonia solution.

The same crystal was used many times, etching and chemical polishing being repeated after each evaporation.

5. Results

Table 1. *Substrate temperatures required to produce epitaxial condensation*

*Fig.**	*Substrate*	*Temp.* (° c)	*Metal*
—	LiF {100} cleaved	600–650	Au
2(*a*)	LiF {100} cleaved	570–620	Ag
—	LiF {100} cleaved	600–650	Cu
—	LiF {100} cleaved	700–750	Pd
—	LiF {110} ground and polished	670–720	Cu
2(*b*)	LiF {110} ground and polished	670–720	Au
2(*c*)	LiF {112} ground and polished	750–800	Au
—	LiF {111} ground and polished	670–720	Au
2(*d*)	Mica {0001} cleaved	270–320	Au
—	Mica {0001} cleaved	270–320	Ag
—	Mica {0001} cleaved	270–320	Cu
—	Mica {0001} cleaved	450–500	Pd

* Figure showing diffraction pattern.

In Table 1 a summary is given of the substrate temperature

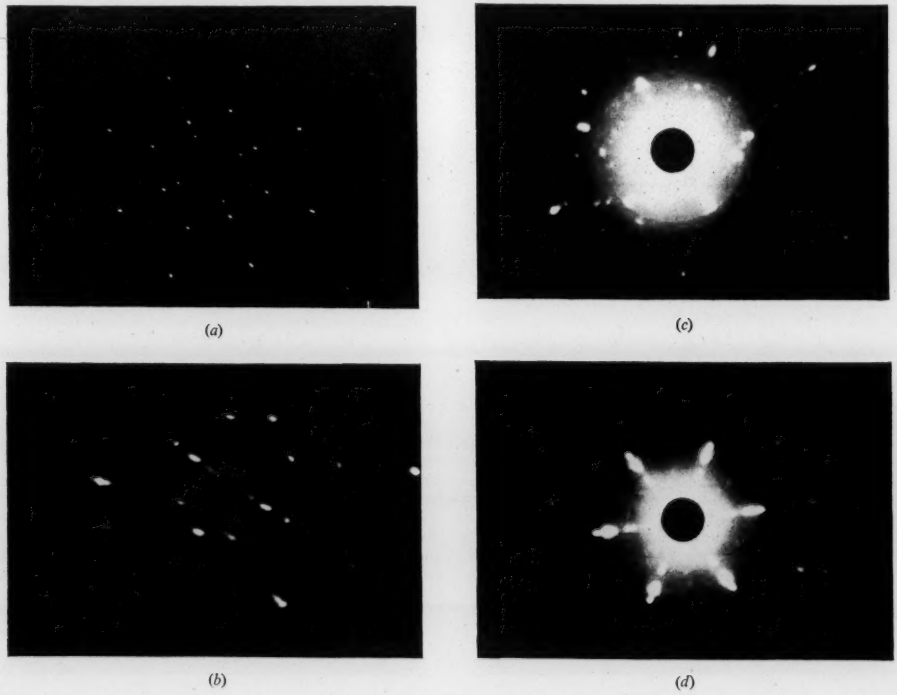


Fig. 2. Back-reflection Laue diffraction patterns from various foils. Obtained with the x-ray beam incident perpendicular to the foil. (*a*) Ag on LiF {100} at 600° c; (*b*) Au on LiF {110} at 690° c; (*c*) Au on LiF {112} at 775° c; (*d*) Au on mica {0001} at 300° c.

necessary to obtain epitaxy for various metal–substrate combinations. In each case foils had such pronounced orientations that their x-ray diffraction patterns resembled those of a single crystal. In Fig. 2 several such patterns are shown, obtained by the Laue back-reflection technique using copper radiation incident perpendicular to the foil and a film to sample spacing of 3 cm.

That these foils are not monocrystalline is shown by the patterns from Au on mica {0001} which clearly exhibits six-fold symmetry, an impossibility for a single cubic crystal (Fig. 2(*d*)). The implication is that two principal orientations are present in an equal amount, each with their {111} planes parallel to the substrate but rotated relative to one another by 180° about their parallel [111] axes. Since the [111] axis is the twinning axis of face-centred cubic metals, these are therefore twin orientations by definition. Diffraction patterns with the beam incident along gold [111] axes other than the surface normal also showed six-fold symmetry. This indicates that twins of the two principal orientations were also present. The presence of twinned regions was confirmed by the photomicrograph of Fig. 4 which shows the surface of a gold foil after etching by ion bombardment.

Foils condensed on LiF {100} show a four-fold symmetry about the substrate normal (Fig. 2(*a*)). When examined along the substrate [111] axis they show a six-fold symmetry, indicating once again that twin orientations are present (Fig. 3) The foils may be described as containing a principal orientation in exact alignment with the substrate crystal with metal {100} parallel to LiF {100} and metal [100] parallel to LiF [100]. In addition they contain four subsidiary orientations which are twins of the principal about its four {111} planes.

A similar state of affairs exists in foils condensed on LiF {110}, where six-fold symmetry is observed about the LiF [111] axis. A principal orientation with metal {110} parallel to LiF {110} and metal [110] parallel to LiF [100] co-exists with its four twin orientations.

The effect of changing the metal on a given substrate is shown in Table 1, the metals copper, silver and gold behave similarly except that silver tends to condense epitaxially at a slightly lower temperature. Palladium requires a temperature higher by 100 deg c in the case of lithium fluoride and 180 deg c for mica. A comparison of the substrates indicates that high-index planes of lithium fluoride require a higher condensation temperature and epitaxy on mica occurs at much lower temperatures than on lithium fluoride.

Fig. 3. Pattern obtained with the beam incident along the principal {111} axis. The apparent six-fold symmetry indicates the presence of twin orientations in Au on LiF {100} at 625° c (lower edge masked by specimen holder).

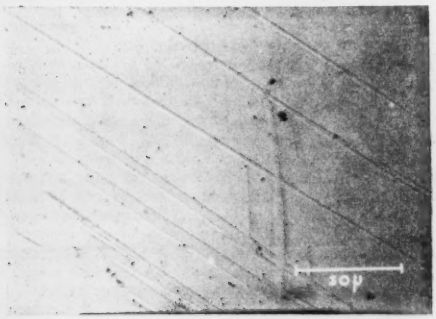


Fig. 4. A photomicrograph of a Au on mica {0001} foil after etching by ion bombardment. The twin boundaries appear as straight lines.

6. Discussion

Current theories (see Pashley 1956) suggest that when epitaxy occurs condensed atoms are initially mobile and may also evaporate from the surface; both the effects being enhanced by increased temperature. In the epitaxial temperature range it is thought that the mobility is high enough to allow ordering whilst the evaporation rate from the surface is small enough to permit growth nuclei to form. It has recently been shown in two independent investigations with the electron microscope that in the first stages of epitaxy, condensation does occur on such nuclei (Bassett and Pashley 1959, Matthews 1959). Bassett and Pashley observed some $2{\cdot}5 \times 10^{11}$ nuclei/cm^2 in the case of Au on NaCl {100} at 300° c, and these were related to surface steps on the substrate. As the mean film thickness increased some nuclei disappeared whilst others expanded and coalesced forming progressively larger islands which eventually merged into a continuous layer. In the case quoted the film did not become parallel sided until a thickness of 800 Å was reached. Electron diffraction showed the films to be substantially monocrystalline and well aligned with the substrate although some twinning was present. Further examination by electron microscopy revealed regions of twin orientations. In both investigations large areas of stacking fault were observed which indicates the rather imperfect nature of crystal growth by epitaxy.

Our own observation that twin orientations are present in thick foils grown on mica may be explained from the reasonable assumption that two equally probable types of nucleus, having a twin relationship, form on a substrate plane which has hexagonal symmetry (Fig. 5). One would then expect the resultant foil to contain a mixture of these two orientations and to exhibit a six-fold symmetry about its normal axis in diffraction.

The origin of twins in the foils grown on a cubic substrate crystal is perhaps due to imperfect condensation on {111} facets of nuclei; the top layer of atoms on such facets having hexagonal symmetry. During the growth of relatively thick foils the direct influence of the substrate lattice on the condensing atoms must disappear and it may be then easier for twins to develop in this way.

The temperatures required for epitaxy reported here are somewhat higher than those found for thin films by other

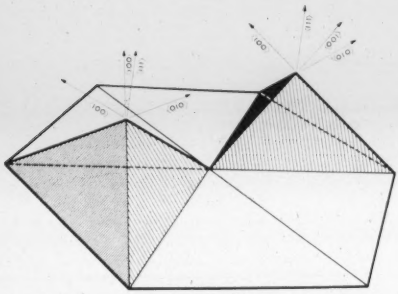


Fig. 5. Showing schematically how twin cubic nuclei may be aligned on a substrate with hexagonal symmetry.

investigators. This difference may arise from the more rapid deposition of metal atoms which would presumably require a higher substrate temperature to induce sufficient atomic mobility. A comparison of the temperatures of epitaxy for various metals on the same substrate (Table 1) and the data in Table 2 suggest that although substrate crystal symmetry exerts a great influence, correlation between lattice constants in substrate and metal is relatively unimportant, and that the temperature of epitaxy bears some relation to the metal's melting point. For instance, copper and gold have considerably different lattice constants, but similar melting points and both exhibit epitaxy at the same temperature on a given substrate. Silver, however, has a slightly lower melting point than gold with a similar lattice and has a slightly lower temperature of epitaxy. Palladium, with an intermediate lattice, but a much higher melting point requires relatively high temperatures for epitaxy. A reasonable inference would be that the mobility of a metal atom on a given substrate depends on the melting temperature of the metal.

Table 2. *Substrate and metal lattice constants*

	Melting point °C	*Lattice constant* Å
Cu	1083	3·61
Ag	961	4·08
Au	1063	4·07
Pd	1555	3·88
LiF	—	4·01
Mica (hex. plane)	—	5·20

References

BASSETT, G. A., and PASHLEY, D. W., 1959, *J. Inst. Metals*, **4**, 449.

MATTHEWS, J. W., 1959, *Phil. Mag.*, **4**, 1017.

PASHLEY, D. W., 1956, *Advances in Physics*, **5**, 173.

THOMPSON, M. W. and NELSON, R. S., 1960, *Proc. Roy. Soc. A*, **259**, 458.

Study of the performance of dielectric thin film beam dividing systems

by L. A. CATALAN,* D.I.C., and T. PUTNER, A.Inst.P., Vacuum Deposition Research Division, Edwards High Vacuum Ltd., Crawley, Sussex

MS. received 12*th April* 1961, *in revised form* 10*th May* 1961

Abstract

A theoretical and practical study has been made to compare the optical performances of single layer $\lambda/4$ TiO_2 and double layer $\lambda/4$ MgF_2–$\lambda/4$ TiO_2 coatings when used as beam-dividing systems. Theoretical and experimental values of the spectral response for light arriving at normal incidence and at 45° incidence are given together with the values for reflectance as a function of angle of incidence for monochromatic light. The optical characteristics of glass–film–glass beam dividing systems are also investigated. It is further shown that the efficiency of a system consisting of a reflecting single $\lambda/4$ layer between glass boundaries is improved by the introduction of a second $\lambda/4$ low refractive index film layer.

1. Introduction

The use of optical interference films of high refractive index in beam-dividing systems is well known and widely referred to in the literature (Holland 1960, Heavens 1955). Such systems have a high optical efficiency due to the absence of light absorption and for this reason are superior to the conventional semi-reflecting coatings using metal films.

The properties (e.g. refractive index, durability and stability) which determine the choice of film material to give the best optical performance are fulfilled fairly well by titanium dioxide as shown by Hass (1952) in his investigations into the preparation and properties of this material. However, it was observed in our laboratory that titanium films may react with the glass substrate when undergoing heat oxidation and barrier layers were used to prevent this (Holland *et al.* 1957). The barrier layers consisted of quarter wavelength films of either aluminium oxide or magnesium fluoride and it was shown that base layers of magnesium fluoride improved the optical performance of TiO_2 beam splitters.

It is the purpose of this paper to extend the previous work by investigating the theoretical and practical performance of the following beam-dividing systems: (i) a single quarter wavelength coating of TiO_2; (ii) a double layer system consisting of quarter wavelength films of MgF_2 and TiO_2 deposited in sequence; (iii) film systems (i) and (ii) cemented between glass media. The spectral response of the systems for visible light is investigated both at normal and 45° incidence.

* Now at the University of Chile, Institute of Physics and Mathematics, Santiago, Chile.

2. Theory

A general expression for the reflectance R of a two-layer coating is (Catalan 1960).

$$R = \left(\frac{\mu_3 E_3 - H_3}{\mu_3 E_3 + H_3}\right)^2 \tag{1}$$

with

$$\mu_3 E_3 \pm H_3 = \left\{(\mu_3 \pm \mu_0)\cos g_1 \cos g_2 \mp \left(\frac{\mu_2}{\mu_1}\mu_0 \pm \frac{\mu_1}{\mu_2}\mu_3\right)\sin g_1 \sin g_2\right\} E_0 + i\left\{\left(\frac{\mu_0}{\mu_1}\mu_3 \pm \mu_1\right)\sin g_1 \cos g_2 + \left(\frac{\mu_0}{\mu_2}\mu_3 \pm \mu_2\right)\sin g_2 \cos g_1\right\} E_0$$

where E is the electric field strength, H is the magnetic field strength

$$\mu_j = n_j/\cos\phi_j$$

or

$$\mu_j = n_j \cos\phi_j \tag{2}$$

according to whether the electrical vector is vibrating perpendicular or parallel to the plane of incidence, n_j being the refractive index of medium j and ϕ_j the angle of incidence at the $j/(j-1)$ boundary (see Fig. 1).

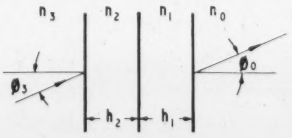


Fig. 1. Convention for the passage of rays through a multi-layer film system.

$$g_j = \frac{2\pi}{\lambda_0} n_j h_j \cos\phi_j \tag{3}$$

is the phase angle corresponding to the vacuum wavelength λ_0 and the optical thickness $n_j h_j \cos\phi_j$ of the layer j.

Similarly, the general expression for the reflectance of a single layer coating is

$$R = \frac{(\mu_2 - \mu_0)^2 \cos^2 g_1 + (\mu_2\mu_0/\mu_1 - \mu_1)^2 \sin^2 g_1}{(\mu_2 + \mu_0)^2 \cos^2 g_1 + (\mu_2\mu_0/\mu_1 + \mu_1)^2 \sin^2 g_1}. \tag{4}$$

When only quarter wavelength coatings are used ($g_J = \pi/2$), the expressions (1) and (4) can be written

$$R = \left(\frac{\mu_2^2\mu_0 - \mu_1^2\mu_3}{\mu_2^2\mu_0 + \mu_1^2\mu_3}\right)^2 \tag{5}$$

for a two quarter-wavelength coating, and

$$R = \left(\frac{\mu_0\mu_2 - \mu_1^2}{\mu_0\mu_2 + \mu_1^2}\right)^2 \tag{6}$$

for a single $\lambda/4$ coating.

It should be pointed out that Eqns (5) and (6) can be applied to any angle of incidence provided Eqn (2) is taken into account. On the other hand, for computations involving $\lambda \neq \lambda_0$, Eqns (1) and (4) have to be employed and the actual phase angle g_λ is given by $g_\lambda = \frac{1}{2}\pi\lambda_0/\lambda$.

3. Experimental

The coatings were deposited on to 2 mm thick soda-lime glass (the refractive index of which was taken as 1·5). The glass surfaces were cleaned by washing in Teepol and water and polishing with finely divided chalk and *iso*-propyl alcohol; finally the glass was exposed to a gas discharge for 10 minutes. The titanium dioxide coatings were always prepared by the heat oxidation in air of thermally evaporated titanium metal as shown by Hass (1952) and Holland *et al.* (1953). The titanium metal was deposited to give a film having a transmission of light of 7·5% at 0·530 mμ which gave a $\lambda/4$ film for $\lambda = 535$ mμ after oxidation in air at 400° C.

The optical measurements were made on a spectrometer fitted with a modified modulated beam photometer as described by Steckelmacher *et al.* (1959). The amplifier used in this work was an improved version of that described in the paper referred to, having zero backing off facilities, and greater sensitivity. The wavelengths were selected using a range of 'Wratten' Kodak absorption filters. The measurements have an accuracy to approximately $\pm 1\%$ of the full-scale deflection when full-scale deflection corresponds to 0–100%, assuming the amplifier to be linear. Where measurements are recorded as being for normal incidence, in practice the angle of incidence fixed by the apparatus was about 15°. The measured values for reflectance include the reflection from the back surface of the glass; it was therefore necessary to make a correction for this in each case.

3.1. *Single $\lambda/4$ layer reflecting system*

The refractive index of an ideal beam splitter for equal transmission and reflection using substrate glasses of refractive index 1·5, should be 2·9 as can be shown by using Eqn (6). The refractive index of heat oxidized TiO_2 films is about 2·7 at $\lambda = 535$ mμ and this appears to be the closest one can get to the ideal index with absorption free metal oxide films.

A titanium dioxide beam splitter was prepared as described above and its reflectance plotted against angle of incidence at $\lambda = 535$ mμ is given in Fig. 2. It can be seen that for unpolarized incident light the reflectance changes by a small amount from 0 to at least 60°. The effects of using a polarized incident beam (i.e. perpendicular R_s and parallel R_p to the plane of incidence) are also shown in Fig. 2. The agreement between measured reflectances and theoretical values is fair. It should be pointed out that for the purposes of calculation the refractive index value of 2·7 was used for the TiO_2 film, thus the small differences between the curves could easily have been due to small index deviations.

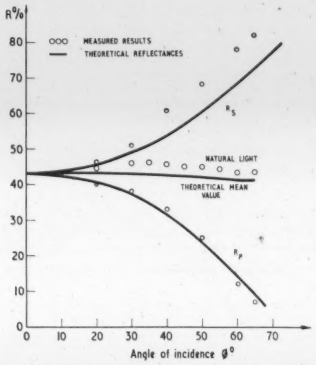


Fig. 2. Reflectance of a $\lambda/4$ TiO_2 coating deposited on to glass ($n_0 = 1{\cdot}50$) as function of the angle of incidence for natural and polarized light ($\lambda_0 = 535$ mμ).

Figure 3 shows measured and theoretical reflectance curves for the spectral response at normal incidence of a $\lambda/4$ TiO_2 layer deposited on to a glass substrate ($n_0 = 1{\cdot}5$). Dispersion is generally neglected when calculating the response of thin film systems, because of the lack of reliable data on the dispersion properties of film materials. However, one set of calculations was made correcting for dispersion using data

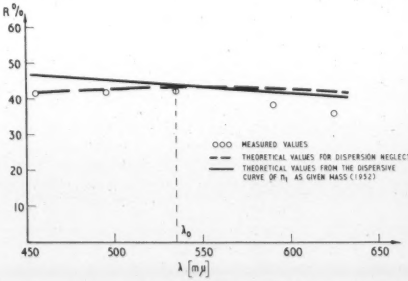


Fig. 3. Spectral reflectance at normal incidence of a $\lambda_0/4$ TiO_2 coating deposited on to glass ($n_0 = 1{\cdot}50$).

for TiO_2 films reported by Hass (1952) and these results are given in Fig. 3, together with those where this property was neglected. The figure shows that though agreement with the measured curve is fair it was not much improved by making the dispersion correction. However, for the remaining film systems containing MgF_2 films dispersion was neglected because data were not available for baked Mg films as used here. The small discrepancies observed in the red region of the spectrum may be caused by an increasing absorption at those wavelengths. These curves show that $\lambda/4$ TiO_2 coatings approach closely the ideal flat response required of a beam-splitter. The theoretical and measured reflectances in Fig. 4 for light incident at 45° also show good agreement and flat response curves.

It can be shown that if the squares of all the incident electric vibrations taken in all directions (where the square

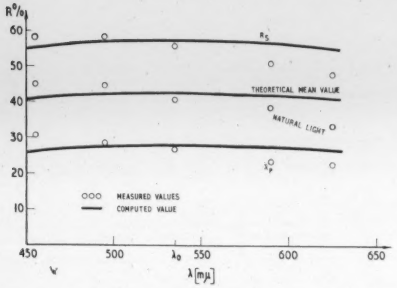


Fig. 4. Spectral response for reflectances at 45° incidence of a $\lambda/4$ TiO_2 coating deposited on to glass ($n_0 = 1{\cdot}50$).

of the amplitude is proportional to energy) are summed for any given angle of incidence, the result is as if the vibrations are always propagated at an angle of 45° to the plane of incidence (the azimuthal angle). Born and Wolf (1959) give a theoretical treatment of this subject. They state: "The reflected light is then said to be partially polarized and its 'degree of polarization' P may be defined to be

$$P = \frac{|R_{||} - R_{\perp}|}{|R_{||} + R_{\perp}|}.$$

Reflectivity $\bar{R}$ is now given by

$$\bar{R} = \frac{\bar{J}^{(r)}}{\bar{J}^{(i)}} = \frac{\bar{J}_{||}^{(r)} + \bar{J}_{\perp}^{(r)}}{\bar{J}^{(i)}} = \tfrac{1}{2}(R_{||} + R_{\perp})$$

where $\bar{J}^{(i)}$ = incident wave energy, $\bar{J}^{(r)}$ = reflected wave energy."

Also given in Fig. 4 are measured and theoretical values for the reflected polarized light (R_s) and (R_p). The theoretical mean curve was determined by summing the mean reflection coefficient values $R_p/2$ and $R_s/2$, where in general $R = (E_r/E_i)^2$ using $\bar{R} = \frac{1}{2}(R_p + R_s)$, i.e. the natural light reflectance. Fig. 4 gives a practical verification of the foregoing statement as can be seen by comparing the calculated mean value and the measured curve for natural light.

3.2. *Double-layer quarter wavelength system*

The double-layer semi-reflecting system consisted of a $\lambda/4$ magnesium fluoride coating, deposited by thermal evaporation, followed by a $\lambda/4$ titanium dioxide coating prepared in the way described above. The titanium film was deposited to give 7·5% transmission as used above. It is interesting to note that the magnesium fluoride layer was not affected by the heat oxidation process. Using Eqn (5) it was verified that a coating of two quarter wavelength layers consisting of two materials with refractive indices corresponding to 1·38 and 2·7 respectively has a reflectance of nearly 50%.

It can be seen in Fig. 5, which gives measured values for reflectance against angle of incidence, that again there is reasonable agreement between the characteristics of the practical coating and the theoretical case. Up to angles of incidence in the region of 55° a reflectance of approximately 50% is maintained fairly constant. It was shown both in theory and in practice that the quarter wavelength thickness of the two films need to be controlled for the same wavelength to produce the ideal 50% reflectance case, that is no mismatching of wavelength thickness between the two layers is necessary.

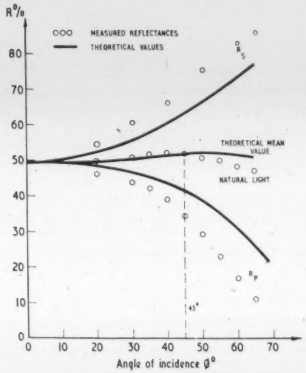


Fig. 5. Reflectance of a two $\lambda_0/4$ layer system of MgF_2+TiO_2 deposited on to glass ($n_0 = 1{\cdot}50$) for $\lambda_0 = 535$ mμ.

A reasonably flat spectral response can be achieved in the range 450 mμ to 650 mμ, see Fig. 6. The theoretical values plotted in this figure were not corrected for dispersion.

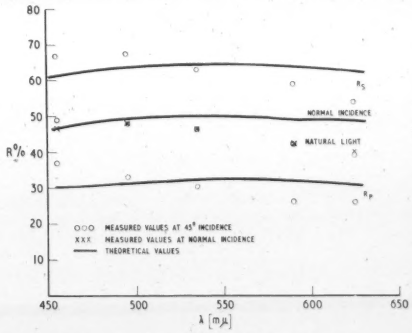


Fig. 6. Spectral response at normal and 45° incidence for natural and polarized light of a two $\lambda_0/4$ layer system of $MgF_2 + TiO_2$ deposited on to glass ($n_0 = 1{\cdot}50$).

By comparing Fig. 6 with Figs 3 and 4 it is clearly demonstrated that both for single and double layer quarter wavelength systems the flat spectral response of the reflectance is very much the same at 45° incidence as it is at normal incidence. Also it is shown that the double-layer system is more efficient than the single-layer coating and its performance approaches remarkably well the ideal beam divider system as described in §1.

3.3. *Beam dividing systems sandwiched between glass media*

A study similar to that made in §3.1 and §3.2 was carried out with single and double quarter wavelength systems sandwiched between glass media ($n_0 = 1{\cdot}5$). The cover glass was cemented to the film system using cellulose caprate which has a refractive index of 1·492. For computation the

refractive index of the cementing medium was taken as being equal to the index of the glass cover. Since in this case the light is incident on an air–glass boundary, for the purposes of calculation we have $\phi_0 \neq \phi_g$, where ϕ_0 is the angle of incidence at the air–glass boundary used in measurement and ϕ_g, used in calculation, is the angle of incidence, due to air–glass refraction, at the glass–film boundary. A correction was made for the front reflection from the cover glass, during calculation. It was shown, as expected from the theory, that in adding the second glass boundary the optical performance of both systems was considerably impaired. The measured reflectances obtained for normal incidence at $\lambda = 535$ mμ were 28% for the single layer system and 34% for the two layer system. It appears then that only slight improvement can be gained by using the double layer system as can be seen by comparing Figs 7 and 9.

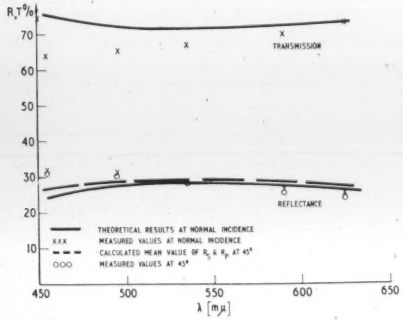


Fig. 7. Spectral response at normal and 45° incidence of a single $\lambda_0/4$ layer TiO_2 sandwiched between glass media ($n_0 = n_3 \simeq 1{\cdot}50$).

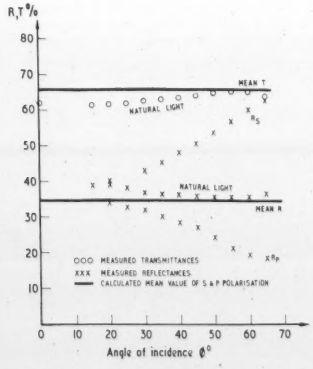


Fig. 8. Reflectance and transmittance of a double layer $MgF_2 + TiO_2$ $\lambda_0/4$ system sandwiched between glass media ($n_0 = n_3 \simeq 1{\cdot}5$) ($\lambda_0 = 535$ mμ).

Conclusion

It has been shown that a double-layer quarter wavelength system consisting of titanium dioxide and magnesium fluoride films gives an improved optical performance compared with the conventional $\lambda/4$ TiO_2 beam divider. In fact, its performance when used between air–low-index glass boundaries approaches the ideal conditions quoted in the introduction. In all the cases considered the measured light reflectance values show remarkably good agreement with the results obtained by using the equations given in the theoretical section. It has been pointed out that when the coatings studied are sandwiched with optical cements between low index glasses, the light reflectance decreases appreciably, the double layer system giving the better optical performance. An investigation of a three-layer quarter wavelength system using a low–high–low index combination might produce a beam divider with more equal reflectance and transmittance values when cemented.

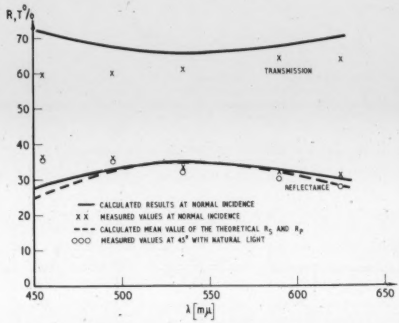


Fig. 9. Spectral response at normal and 45° incidence of a double layer $MgF_2 + TiO_2$ $\lambda_0/4$ system sandwiched between glass media ($n_0 = n_3 \simeq 1{\cdot}50$).

Acknowledgments

Acknowledgments are made to the Directors of Edwards High Vacuum Ltd., for permission to publish this paper and to Mr. L. Holland, Head of the Vacuum Deposition Research Division, for his guidance throughout this work.

References

BORN, M., and WOLF, E., 1959, *Principles of Optics* (London: Pergamon Press).

CATALAN, L. A., 1960, Thesis (University of London), p. 21.

HASS, G., 1952, *Vacuum*, **2**, 331.

HEAVENS, O. S., 1955, *Optical Properties of Solid Films* (London: Butterworths).

HOLLAND, L., 1960, *Vacuum Deposition of Thin Films* (London: Chapman and Hall).

HOLLAND, L., HACKING, K., and PUTNER, T., 1953, *Vacuum*, **3**, 159.

HOLLAND, L., PUTNER, T., and BATEMAN, S., 1957, *J. Opt. Soc. Amer.*, **47**, 668.

STECKELMACHER, W., PARISOT, J. M., HOLLAND, L., and PUTNER, T., 1959, *Vacuum*, **4**, 171.

Counting statistics in x-ray spectroscopy

by R. C. STANLEY, B.Sc., Materials Division, Central Electricity Research Laboratory, Leatherhead, Surrey

MS. received 3rd May 1961

Abstract

An investigation has been made to see whether the secondary x-ray photons arrive in a random manner in a commercially available x-ray fluorescence spectrometer, and if so, how may the available counting time be best used. A sufficiently large number of observations has been taken to allow a full statistical analysis to be made and compared with theoretical predictions. In addition, an experimental investigation has been made into the timing accuracy and its effect on the final count rate.

It was found that for very high counting rates, above 10 000 *counts/s, or long counting times, above about* 2 *minutes, small instrumental errors tend to occur, but for count rates generally obtained in practice, the x-ray photons are emitted in a random manner and a Poisson distribution is obtained. In practice one could expect a single observation to be within* 0·6% *of the true value for a* 256 000 *count or* 1·3% *for a* 64 000 *count.*

The necessary statistics for measurements of a given accuracy are set out in a form readily usable by the x-ray spectroscopist, together with formulae and procedure for obtaining the greatest net accuracy when only a limited time is available.

Introduction

In an x-ray spectrometer, x-rays may be produced in two ways. In x-ray fluorescence analysis, the x-rays from the anode of an x-ray tube are used to generate secondary x-ray photons from a specimen. In the primary x-ray emission method, the x-rays are produced by the direct impingement of electrons on to the specimen. It is tacitly assumed that these x-ray photons, produced in either way, are emitted in a random manner both with respect to direction and time.

Two problems of importance to the spectroscopist then arise:

(i) Is the process in fact random in time? Direct tests of this are difficult, but it is simple to investigate whether counting rates are random in a particular spectrometer. If this is the case, it is unlikely that instrumental factors are responsible and the random process must be the generation of x-ray photons.

(ii) If random, how many counts must be taken to obtain the required degree of accuracy in count rate?

Other authors (Klug and Alexander 1954, Liebhafsky *et al.* 1955, 1960, Mack and Spielberg 1958, Parrish 1956) have discussed the general statistics of counting methods. In this communication it is proposed to give in a readily usable form all the statistics necessary for x-ray spectroscopic measurements of a given accuracy, and to examine the two problems raised with special reference to a commercially available x-ray fluorescence spectrometer.

Statistical definitions

Assume first a random emission process having a constant mean counting rate. Then, since there are a large number of atoms in the sample and the probability of a particular one being excited is small, the distribution approximates to a Poisson distribution. Thus the probability of recording a count rate x for a given number of counts N, when the average rate is $\bar{x}$, is

$$P_N(x) = \frac{(\bar{x})^x}{x!}e^{-\bar{x}}. \tag{1}$$

In the present case $x^{\frac{1}{2}} \gg 1$ and the distribution also approximates to a Gaussian distribution of standard deviation σ,

where $$P_N(x) = (2\pi)^{-\frac{1}{2}}\sigma \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right). \tag{2}$$

We can then make use of the following formulae and definitions.

(i) 'Mean' $\bar{x}$, the average count rate, is the sum of all the count rates determined for equal numbers of counts (or equal times) divided by the number of observations n. If a large number of observations is made with the same precision, this constitutes the most probable approximation to the true value.

(ii) 'Standard deviation' σ, is the root mean square deviation, given by

$$\sigma = \left(\frac{\Sigma(x-\bar{x})^2}{n}\right)^{\frac{1}{2}} \tag{3}$$

where x is the count rate determined in a single observation. This is the maximum deviation from the true count rate that may be expected to occur in 68% of a large number of observations.

Also, for a Poisson distribution

$$\sigma = N^{\frac{1}{2}}/t \tag{4}$$

where N is the total number of counts to be observed, provided that the time t can be observed without error.

(iii) 'Relative standard deviation' Δ. This enables sets of observations to be compared and is equal to the standard deviation divided by the mean. It is conveniently expressed as a percentage. If a large number of measurements be made, it may be expected that at least 68% of the observations will differ from the true value by the relative standard deviation or less.

$$\Delta = \frac{\{\Sigma(x-\bar{x})^2/n\}^{\frac{1}{2}}}{\bar{x}} \times 100\% \tag{5}$$

or for a Poisson distribution

$$\Delta = \frac{N^{\frac{1}{2}}}{N} \times 100\% = \frac{100}{N^{\frac{1}{2}}}\%. \tag{6}$$

Thus if a large number of observations be made on a given number of counts, the relative standard deviation (R.S.D.%) calculated from the root mean square deviation should agree approximately with that calculated as $(100/N^{\frac{1}{2}})\%$. If there is no agreement, the distribution is not a Poisson distribution and it must be assumed that the basic process is not random or that the particular instrument is at fault as a result, for example, of the occurrence of spurious counts.

In addition, two other methods of expressing the error are sometimes used. These are:

(i) 'Reliable error' R—a value such that 90% of the readings will have errors less than this value. It is equal to $1{\cdot}64 \times$ R.S.D., and is usually expressed as a percentage, so that

$$R = \frac{164}{N^{\frac{1}{2}}}\%. \tag{7}$$

(ii) 'Probable error' P—that error for which there is an equal chance of it being exceeded as not. It is the maximum deviation from the true value that may be expected in 50% of a large series of measurements and is equal to $0{\cdot}68 \times$R.S.D. and again is usually expressed as a percentage, so that

$$P = \frac{68}{N^{\frac{1}{2}}}\%. \tag{8}$$

As an example, for a count of 16000, one could expect a standard deviation of $(16\,000)^{\frac{1}{2}} = 126{\cdot}5$, which corresponds to a relative standard deviation of $100/(16\,000)^{\frac{1}{2}} = 0{\cdot}8\%$. In addition, the reliable error would be $1{\cdot}64 \times 0{\cdot}8 = 1{\cdot}3\%$ and the probable error $0{\cdot}68 \times 0{\cdot}8 = 0{\cdot}5\%$.

Randomness of count rate

To check the randomness of the arrival of the x-ray photons, the following investigations were made.

One hundred successive observations were made of the time required to accumulate a given number of counts. These were obtained using a Philips, PW1010/30 stabilized x-ray generator, with a PW1050/25 goniometer, PW1540 vacuum spectrograph, and a PW1051/30 electronic circuit panel. The counts were made on the first order copper K_{α_1} peak ($2\theta = 44{\cdot}96$, LiF analysing crystal) without pulse height discrimination, using a PW1964/10 scintillation counter (850 v) and a tungsten anode x-ray tube. This was operated at 6 mA and sets of observations were made at various kilovoltages, to obtain different counting rates. In the Philips' instrument, the time required to accumulate a predetermined number of counts is obtained. Measurements of time were made to an accuracy of 0·05 second on the instrument's clock.

A check was also made on the accuracy of the clock. In the Philips' instrument, there are facilities for counting on a test signal (100 counts/s) derived from the 50 c/s mains frequency. Sets of one hundred successive observations were made on this signal for various total counts, to give different counting periods. A continuous check was also made on the mains frequency against a quartz crystal oscillator with a claimed accuracy to better than 0·01%. Each observed time was then corrected accordingly. The results are given in Table 1, assuming a Gaussian distribution.

From these it is seen that there is a constant standard deviation of 0·04 s independent of the time being recorded. It would thus appear that all the errors due to the clock occur during starting and stopping and not in the running period.

Table 1. *Observed deviations for clock*

Recorded time	*S.D.* σ	*R.S.D.* Δ
20 s	0·042 s	0·21%
40 s	0·040 s	0·10%
80 s	0·039 s	0·05%
160 s	0·040 s	0·025%

The results for count rates obtained for various kilovoltages are summarized in Table 2. For the lower count rates, the time to accumulate a hundred observations became long and there was a long-term drift (up to 4% in 6 hours). This was possibly due to temperature changes in the surroundings. In the practical use of the spectrometer, measurements on a sample would be completed in a few minutes, and usually by ratio with a known standard. Thus the effects of any long-term drift in the spectrometer would not be evident and must therefore be compensated for in the present series of observations. This was done by fitting a plot of count rate against order of observation by the method of least squares and computing the deviations from this curve.

In the sixth column of the table is given the maximum error obtained in the individual counts, expressed as a percentage error of the mean, and in the last column is given the number of times the 99% confidence limit is exceeded in a set of 100 readings. For a 99% confidence limit, one would expect only 1% of the observations to have an error greater than 0·59% of the mean for measurements of 256000 counts, or greater than 1·19% for measurements of 64000 counts.

Table 2. *Observed deviations for* (*a*) 256000 and (*b*) 64000 *counts* (100 *observations*)

(*a*) 256000 counts.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
11	1213	211·0	3·4	0·28	1·05	5
12	2558	100·1	5·4	0·21	0·57	0
13	4782	53·5	7·1	0·15	0·44	0
14	7246	35·3	14·7	0·20	0·51	0
15	10067	25·4	25·6	0·25	0·71 (2×)	2
16	13098	19·5	34·8	0·27	0·74	1

(*b*) 64000 counts.

10	393	162·8	1·6	0·41	1·14	0
11	1304	49·1	4·8	0·37	0·92	0
12	2580	24·8	11·4	0·44	1·28	2
13	4858	13·2	20·5	0·42	0·95	0

(1) Voltage (kv); (2) mean counts/s $\bar{x}$; (3) mean time for one observation t (s); (4) S.D. σ (counts/s); (5) R.S.D. Δ(%); (6) maximum error as % of $\bar{x}$; (7) % of observations > 99% confidence limit.

In Table 2(*a*), the results are given for counts of 256000. With a Poisson distribution these should give an R.S.D. of $100/N^{\frac{1}{2}} = 0{\cdot}20\%$. Table 2(*b*) gives similar results for counts of 64000. The R.S.D. should now be 0·40%.

As can be seen, the results are in good agreement with the experimental R.S.D.'s computed from $(100\sigma/\bar{x})\%$. Thus a Poisson distribution for the count rates obtained by the spectrometer may be assumed.

In practice, one should be able to take a reading and expect that in 99 cases out of 100 the results will be within 0·6% of the true value for a 256000 count or within 1·3% for a 64000 count. This, however, does not appear to be

true for the longest counting time and possibly for the fastest counting rates. For the longest counting time, it is presumed that here there is more chance of external interference due to mains fluctuations, for example. Thus in practice, counting times should not be longer than about 2 minutes, if this can conveniently be arranged. Also, with the highest counting rate, there is an increase in the R.S.D., presumably due to partial paralysis in the electronic counting circuits. Count rates should therefore be kept below about 10000 counts/s.

Errors due to deviations of the clock can be shown to be negligible. The corrected standard deviation of the count rate is $\sigma_{\bar{x}}/t$, where $\sigma_{\bar{x}}$, the standard deviation of the number of counts, is given by (Parrish 1956)

$$\sigma_{\bar{x}/t} = \frac{\bar{x}}{t}\left(\frac{\sigma_{\bar{x}}^2}{\bar{x}^2} + \frac{\sigma_t^2}{t^2}\right)^{\frac{1}{2}} \tag{9}$$

where $\sigma_{x/t}$ and σ_t are the standard deviations of the observed count rate and the clock reading (i.e. 0·04 s), and t is the mean time to observe a mean count $\bar{x}$.

Thus the greatest correction to the observed count rate would be the observation involving the least time. For example, the R.S.D.% of the count rate in the 64000, 13 kv observation, would be reduced from 0·422 to 0·421 with error-free timing. That is, the standard deviations of the observed count rates are not materially affected by the timing and must be presumed to be entirely due to the randomness of the emitted x-ray photons.

Counts for required accuracy

With a random arrival of x-ray photons, the number of counts required to obtain a given accuracy may be calculated. An error of 1Δ means, in this instance, an error of 1 R.S.D. expressed as a percentage. For example, the R.S.D. of 64000 counts would be $100/(64000)^{\frac{1}{2}}$, that is, 0·4%, and an error of 1Δ would mean that variations of 0·4%, or less, from the true value would be expected to occur in 68% of a large series of measurements. Similarly, an error of 2Δ would mean variations of 0·8% or less would be expected in 96% of the measurements.

Table 3 gives the expected errors calculated in this way.

Table 3

N	1Δ(%)	2Δ(%)	3Δ(%)
1000	3·16	6·32	9·49
2000	2·24	4·47	6·71
4000	1·58	3·16	4·74
8000	1·12	2·24	3·35
16000	0·79	1·58	2·37
32000	0·56	1·12	1·68
64000	0·40	0·79	1·19
128000	0·28	0·56	0·84
256000	0·20	0·40	0·59
Probability	68%	96%	99%

Thus, to obtain the required accuracy, the appropriate number of counts to be taken may be found from the table, or calculated as follows. For example, suppose an accuracy of 1% or better is required in 99% of the determinations. That is, a 3Δ variation (see Table 3) of 1% is required. Divide by 3 to convert to a 1Δ error of 0·33%. Then from Eqn (6) $N = 92809$, i.e. about 93000 counts must be taken to achieve the required accuracy.

However, since the calculated accuracy can only be the probable accuracy, there must still exist an uncertainty if only one measurement is made. In addition, there is still the possibility of some instrumental failure, or error in the counting due to sudden large voltage fluctuations in the public power supply.

There is, consequently, much to be said for taking several measurements of fewer counts than that required for statistical reasons and summing. For example, on a total of N counts the standard deviation would be $N^{\frac{1}{2}}$. If, however, three counts of $\frac{1}{3}N$ were made, each of standard deviation $(\frac{1}{3}N)^{\frac{1}{2}}$, then, on summing the counts for a total N, the total standard deviation would be $(\frac{1}{3}N + \frac{1}{3}N + \frac{1}{3}N)^{\frac{1}{2}} = N^{\frac{1}{2}}$, as before. That is, the same statistical accuracy is achieved by obtaining the required total counts by summing a series of lesser counts. This method of counting has the advantage that it is at once obvious if there is any marked drift, or any marked error in one of the measurements. If the measurements are in close agreement, say, within 1% of each other, confidence is established in the results. From Table 2 it is seen that errors of this order should be easily attained in practice. If, in addition, the measurements on the specimen are alternated with measurements on a stable standard, results may be compared by ratio and the effects of any long-term drift in instrumental performance nullified.

Effect of 'background'

So far, the effect of background has been neglected, but in practice this is rarely possible and it is necessary to make measurements on the peak height and on the background, and find the difference.

Then if t_p and t_b are the times obtained to accumulate the same number of counts N on the peak and background respectively, the net counting rate will be $N/t_p - N/t_b$ with an S.D. of $N^{\frac{1}{2}}/t_p$ for the peak measurement and $N^{\frac{1}{2}}/t_b$ for the background. Then the net S.D. is given by

$$\left\{\left(\frac{N^{\frac{1}{2}}}{t_p}\right)^2 + \left(\frac{N^{\frac{1}{2}}}{t_b}\right)^2\right\}^{\frac{1}{2}} \tag{10}$$

and the R.S.D. is

$$\left\{\left(\frac{N^{\frac{1}{2}}}{t_p}\right)^2 + \left(\frac{N^{\frac{1}{2}}}{t_b}\right)^2\right\}^{\frac{1}{2}} \Big/ \left(\frac{N}{t_p} - \frac{N}{t_b}\right) \tag{11}$$

or

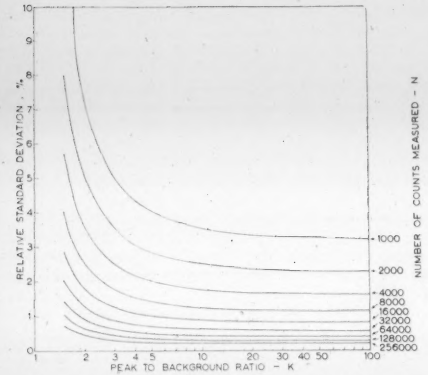
$$100\left(1 + \frac{1}{K^2}\right)^{\frac{1}{2}} \Big/ N^{\frac{1}{2}}\left(1 - \frac{1}{K}\right)\% \tag{12}$$

where $K = t_b/t_p$. The Figure shows plots of the relative standard deviation (R.S.D.%) for various values of K, the peak to background ratio, and of N, the number of counts measured. As can be seen, for peak–background ratios in excess of 10 the effect of background on the R.S.D. can be neglected without serious error.

When the background is of importance, the problem of the relative distribution of counting time between the peak and the background becomes of interest. Only a limited time may be available for a particular investigation, and the problem is then to determine the number of counts, N_p and N_b, on the peak and background respectively, to obtain the required net R.S.D., Δ_D. If Δ_p and Δ_b be the corresponding R.S.D.'s, then

$$\Delta_D = \frac{\{(x_p\Delta_p)^2 + (x_b\Delta_b)^2\}^{\frac{1}{2}}}{(x_p - x_b)} \times 100\% \tag{13}$$

where x_p, x_b are the approximate count rates obtained on



Various values of K, the peak to background ratio, and N, the number of counts measured, necessary to obtain a given R.S.D. (%).

the peak and background respectively, and $(x_p - x_b)$ is the net counting rate. Also,

$$\Delta_p = \frac{1}{N_p^{\frac{1}{2}}} = \frac{1}{(x_p t_b)^{\frac{1}{2}}} \tag{14}$$

where t_p is the time required to obtain N_p counts on the peak. Similarly, for the background,

$$\Delta_b = \frac{1}{N_b^{\frac{1}{2}}} = \frac{1}{(x_b t_b)^{\frac{1}{2}}} \tag{15}$$

therefore

$$\Delta_D = \frac{100}{x_p - x_b}\left(\frac{x_p}{t_p} + \frac{x_b}{t_b}\right)^{\frac{1}{2}} \%. \tag{16}$$

Putting $K = t_b/t_p$ and $T = t_b + t_p$, we have

$$\Delta_D^2 = \frac{10^4}{T(x_p - x_b)^2}\left(x_p K + x_b + x_p + \frac{x_b}{K}\right). \tag{17}$$

Differentiating, then for a minimum,

we find

$$K = \left(\frac{x_b}{x_p}\right)^{\frac{1}{2}}. \tag{18}$$

Combining this with Eqn (16) and by subsequent manipulation, it is found that the number of counts N_p to be taken on the peak to obtain a given net R.S.D., $\Delta_D\%$, is given by

$$N_p = \frac{10^4}{\Delta_D^2}\frac{x_p\{x_p + (x_p x_b)^{\frac{1}{2}}\}}{(x_p - x_b)^2}. \tag{19}$$

Similarly,

$$N_b = \frac{10^4}{\Delta_D^2}\frac{x_b\{x_b + (x_p x_b)^{\frac{1}{2}}\}}{(x_p - x_b)^2}. \tag{20}$$

Acknowledgments

Thanks are due to D. R. Holmes and G. Spurr of these Laboratories for helpful discussions, to R. J. Lewis who made the measurements, and to Dr. J. S. Forrest, Director of Laboratories, for permission to publish.

References

COOK, G. B., and DUNCAN, J. F., 1952, *Modern Radiochemical Practice* (Oxford: Clarendon Press), p. 60.

KLUG, H. P., and ALEXANDER, L. E., 1954, *X-ray Diffraction Procedures* (New York: Wiley), p. 271.

LIEBHAFSKY, H. A., PFIEFFER, H. G., and ZEMANY, P. D., 1955, *Analyt. Chem.*, **27**, 1257.

—— 1960, *Proceedings of the Second International Symposium on X-Ray Microscopy and X-Ray Microanalysis, Stockholm*, 1960. Ed. A. Engstrom, V. Cosslett and H. Pattee (London: Van Nostrand), p. 321.

MACK, M., and SPIELBERG, N., 1958, *Spectrochim. Acta*, **12**, 169.

PARRISH, W., 1956, *Philips Tech. Rev.*, **17**, 206.

Measurement of electrical resistivity by a mutual inductance method

by R. G. CHAMBERS, M.A., Ph.D., and J. G. PARK,* M.A., D.Phil., H. H. Wills Physics Laboratory, Bristol

MS. received 1st *May* 1961

Abstract

The resistivity of a sample can be deduced from the change in mutual inductance between two coils when the sample is inserted. It is shown that with simple equipment for measuring mutual inductance over a range of frequencies, the method can be used to measure resistivities from 2×10^{-9} *ohm cm upwards, and the necessary functions are tabulated.*

Introduction

THE self-inductance L of a coil, or the mutual inductance M between a pair of coils, is reduced when a non-ferromagnetic conductor is placed in the magnetic field generated by alternating current, and the losses are at the same time increased. We may represent the decrease in L or M on bringing up the conductor by $\delta L = \delta L' + i\delta L''$ or $\delta M = \delta M' + i\delta M''$. By measuring δL or δM at a number of frequencies it is possible, knowing the geometry of the system, to deduce the resistivity of the conductor. Using simple equipment, the accuracy of such measurements is limited to about 1%, but the method has the advantage that no direct electrical connections need to be made to the sample. The method has been described previously by a number of authors (Kouwenhoven and Daiger 1934, Grube and Speidel 1940, Laurmann and Shoenberg 1949, Fraser and Shoenberg 1949, Rorschach and Herlin 1951, Van den Berg and Van der Marel 1954); the object of the present paper is to discuss the interpretation of the results in rather more detail.

Theory

If the sample is isotropic in resistivity, the alternating field $H = H_0 e^{i\omega t}$ inside it satisfies the equation

$$\nabla^2 H = 2iH/\delta^2 \tag{1}$$

subject to the appropriate boundary conditions. Here δ is the skin depth for a plane surface: if ρ is measured in microhm cm, f in kc/s and δ in cm,

$$\delta = \frac{1}{2\pi}\left(\frac{\rho}{f}\right)^{1/2}.$$

Solutions of the related equation $\nabla^2 H = H/\lambda^2$ are given, for instance, by Shoenberg (1952) for an infinite circular cylinder parallel to the applied field, for a thin plate with the field parallel to its surface, and for a sphere. The solutions of (1) follow on replacing λ by $\delta/(2i)^{1/2}$.

A well-defined geometry is most simply achieved by using a mutual inductance with a long primary winding, and a short secondary coil wound over the centre of the primary, and measuring the change δM in the mutual inductance on inserting a long cylindrical specimen. (The primary and secondary may, of course, be interchanged without affecting δM). The flux through the secondary is then not appreciably perturbed by end effects, and the solution of (1) for an infinite circular cylinder may be used. This solution may be expressed in terms of an effective volume susceptibility χ or permeability μ:

$$\chi = \frac{1}{4\pi}\frac{J_2(xi^{3/2})}{J_0(xi^{3/2})}; \quad \mu = \frac{2J_1(xi^{3/2})}{xi^{3/2}J_0(xi^{3/2})}$$

Here the J_ν are Bessel functions of the first kind, and $x = \sqrt{2}a/\delta$, where a is the specimen radius. Thus

$$\rho/a^2 = 8\pi^2 f/x^2 \tag{2}$$

where again ρ is expressed in $\mu\Omega$ cm, f in kc/s and a in cm. If A_c and A_s are the cross-sectional areas of the primary coil and the specimen and the mutual inductance changes from M_c to $M_c - \delta M$ when the specimen is inserted, we have $\delta M = \delta M' + i\delta M'' = -4\pi\chi M_c A_s/A_c$. At high frequencies, where $\delta \ll a$, little flux penetrates into the specimen, χ tends to $-1/4\pi$ and δM tends to the limiting value $\delta M_0 = M_c A_s/A_c$. If the specimen becomes superconducting, $\chi = -1/4\pi$ at all frequencies (assuming that the penetration depth λ is much less than a) and δM_0 can be measured directly. Writing $\delta M'/\delta M_0 = m'$, $\delta M''/\delta M_0 = m''$, we have

$$m' + im'' = -J_2(xi^{3/2})/J_0(xi^{3/2}). \tag{3}$$

The variation of m' and m'' with x is shown in the Table and in Fig. 1; the Table also shows the values of

$$\alpha = x^2 m''/6m'.$$

For small x,

$$m' = \frac{x^4}{48}\left(1 - \frac{57x^4}{1920}\cdots\right), \quad m'' = \frac{x^2}{8}\left(1 - \frac{55x^4}{1920}\cdots\right)$$

and for large x,

$$m' = 1 - \frac{\sqrt{2}}{x} - \frac{\sqrt{2}}{8x^3}\cdots, \quad m'' = \frac{\sqrt{2}}{x} - \frac{1}{x^2} - \frac{\sqrt{2}}{8x^3}\cdots$$

(Fraser and Shoenberg 1949); these series are useful for $x \leqslant 1$ and $x \geqslant 5$ respectively. At $x = 2{\cdot}49_6$, $m' = m''$, and m'' passes through a maximum at $x_m = 2{\cdot}51_8$. The corresponding frequency is given from (2) by

$$f_m = 0{\cdot}0803\rho/a^2. \tag{4}$$

* Now at the Department of Physics, Imperial College, London, S.W.7.

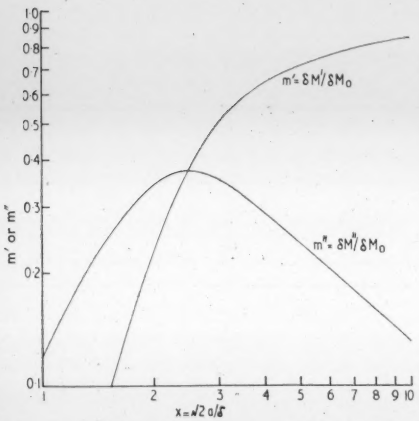


Fig. 1. Variation of m' and m'' with x (x is proportional to the square root of the frequency).

x	m'	m''	α
0·0	0·0000	0·0000	1·0000
0·2	0·0000	0·0050	1·0000
0·4	0·0005	0·0200	1·0000
0·6	0·0027	0·0448	1·0001
0·8	0·0084	0·0791	1·0004
1·0	0·0202	0·1214	1·0010
1·2	0·0407	0·1679	1·0022
1·4	0·0718	0·2205	1·0041
1·6	0·1142	0·2695	1·0069
1·8	0·1667	0·3121	1·0109
2·0	0·2262	0·3449	1·0163
2·2	0·2884	0·3660	1·0235
2·4	0·3495	0·3760	1·0330
2·6	0·4063	0·3768	1·0451
2·8	0·4569	0·3707	1·0601
3·0	0·5009	0·3600	1·0780
3·2	0·5387	0·3469	1·0988
3·4	0·5710	0·3327	1·1226
3·6	0·5987	0·3186	1·1494
3·8	0·6224	0·3050	1·1792
4·0	0·6429	0·2921	1·2118
4·2	0·6609	0·2802	1·2467
4·4	0·6769	0·2693	1·2838
4·6	0·6912	0·2593	1·3229
4·8	0·7041	0·2500	1·3637
5·0	0·7159	0·2415	1·4059
5·2	0·7268	0·2337	1·4493
5·4	0·7370	0·2265	1·4934
5·6	0·7464	0·2196	1·5380
5·8	0·7553	0·2132	1·5826
6·0	0·7635	0·2070	1·6267

The ratio ρ/a^2 can be determined experimentally in three ways:

(*a*) If δM_0 is known, by measuring $\delta M'$ or $\delta M''$ at a single frequency. From the ratio m' or m'', the value of x can be found from the Table, and ρ/a^2 from (2). If $\delta M''$ alone is measured, it is also necessary to know whether the measuring frequency is above or below f_m.

(*b*) By measuring $\delta M'$ or $\delta M''$ over a range of frequencies, and scaling the experimental graph of $\delta M'$ (or $\delta M''$) against $f^{1/2}$ to fit the theoretical graph of m' (or m'') against x. This is most simply done by plotting both graphs logarithmically, as in Fig. 1. Then the vertical scaling factor gives δM_0 (which therefore need not be known beforehand), and the horizontal scaling factor gives $f^{1/2}/x$, and hence ρ/a^2 from (2) or (4).

(*c*) The simplest and most convenient method, which again involves no knowledge of δM_0, is to measure both $\delta M'$ and $\delta M''$ at a single frequency. Then the ratio $\delta M''/\delta M' = m''/m' = 6\alpha/x^2$ is a unique function of x, and an accurate value of x is readily found by successive approximations, using the Table. The value of ρ/a^2 follows as before from (2), and δM_0 can be found from

$$\delta M_0 = \delta M'/m'(x) = \delta M''/m''(x).$$

The analysis is particularly simple at low frequencies ($f \ll f_m$) where $\alpha \simeq 1$. This method has the advantage that residual end-effect errors, which chiefly affect δM_0 (in a frequency-dependent fashion) are virtually eliminated.

If end-effect errors can be neglected, and the coil constant M_c/A_c is known, the value of a^2 can be deduced at once from the measured δM_0 using $\delta M_0 = \pi a^2 M_c/A_c$ so that it is unnecessary to measure a directly. In particular, if method (*b*) is used, we have $m' \simeq m'' = 0{\cdot}3774$ at $f = f_m$ so that from Eqn (4) the impedance changes at f_m give ρ directly:

$$\omega\delta M' \simeq \omega\delta M'' = 0{\cdot}598\rho M_c/A_c \qquad (5)$$

where $\omega\delta M$ is expressed in $m\Omega$, ρ in $\mu\Omega$ cm, and M_c/A_c in μH cm^{-2}.

In practice it is desirable not to work at frequencies much higher than f_m, since at such frequencies the field penetration is small and the result is determined only by the resistivity of the surface layers of the specimen, rather than by the bulk resistivity, and may be affected by surface roughness. Moreover, very pure specimens at low temperatures will exhibit anomalous skin effect behaviour when the electronic mean free path l becomes comparable with δ and under these conditions Eqn (1) breaks down and the present analysis becomes inapplicable. The condition for (1) to be valid is $l/\delta \lesssim 10^{-1}$ as shown by Reuter and Sondheimer (1948). For most metals $\rho l \sim 10^{-5}$ $\mu\Omega$ cm^2 (Chambers 1952) so that this condition can be written

$$f \lesssim 2 \times 10^6 \rho^3 \qquad (6)$$

with f in kc/s, ρ in $\mu\Omega$ cm. Similarly, size-effect errors will arise if l is comparable with the specimen radius, and to avoid these we require $l/a \lesssim 10^{-2}$. Thus, using the above estimate of ρl we find that we need $a \gtrsim 10^{-3}/\rho$. With the aid of Eqn (4) this can be put in a form similar to (6):

$$f_m \lesssim 10^5 \rho^3. \qquad (7)$$

If errors of 5–10% in the determination of ρ can be tolerated, these frequency limits can probably be exceeded by factors of 10 and 100 respectively (corresponding to $l/\delta \lesssim 0{\cdot}3$, $l/a \lesssim 0{\cdot}1$).

Experimental details

As a practical example, we describe briefly the apparatus we have used to measure the residual resistivity of dilute Hg–Cd alloys (Chambers and Park 1960). The samples were 2 cm long and 0·5–1 mm in diameter, and their resistivities ranged between 0·0025 and 0·5 $\mu\Omega$ cm. The 200 μH mutual

inductance M into which the samples were inserted had an effective cross-sectional area A_c of about $0{\cdot}5\ \text{cm}^2$, so that δM_0 was about 1–3 μH. To measure δM, a Hartshorn bridge was used (Fig. 2). From the changes in R and M_b

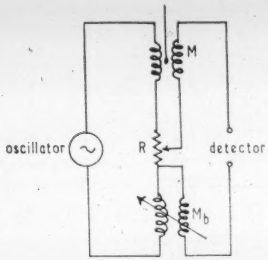


Fig. 2. The Hartshorn bridge.

needed to restore balance on inserting the sample, we have $\delta M' = \delta M_b$, $\delta M'' = \delta R/\omega$. Provision was made for inserting and withdrawing the sample from outside the cryostat, so that δR and δM_b could be measured directly. This is necessary because the empty-coil balance point itself varies with temperature and frequency. The compensating inductance M_b was in three parts: a main fixed compensator, similar to M, inside the cryostat, a small uncalibrated variable for zero-setting, and a carefully calibrated variable with ranges $+0{\cdot}45n$ to $-0{\cdot}45n$ μH, where $n = 1, 3, 10, \ldots 1000$. The variable resistance R was a conventional 0·1 ohm low-inductance slide-wire potentiometer, supplemented by $\times 10^{-1}$ and $\times 10^{-2}$ attenuators when necessary. The bridge was fed through a power amplifier by an audio oscillator (20 c/s–20 kc/s), and the output from the bridge was fed through a conventional audio amplifier to an oscilloscope.

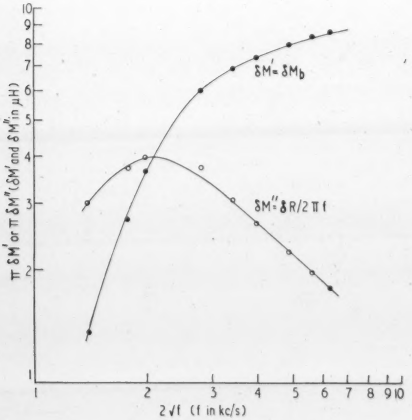


Fig. 3. An example of the use of method (*b*). The theoretical curves that fit the points best are shown superimposed. (The values of $\delta M'$, $\delta M''$ and $\sqrt{f}$ have been multiplied by factors which bring all the points within one cycle of logarithmic graph paper, for convenience in fitting the curves.)

If δM_b is measured in μH, δR in mΩ and f in kc/s, then $m''/m' = \delta R/2\pi f \delta M_b$, and from Eqn (2) we have simply $\rho/a^2 = (2{\cdot}094/\alpha)\delta R/\delta M_b$ for method (*c*) of measurement. Thus in this method it is not necessary to measure the frequency accurately, if one works at low frequencies were $\alpha \simeq 1$. Finally, as an example of the use of method (*b*), Fig. 3 shows typical plots of $\delta M'$ and $\delta M''$ against $f^{1/2}$, with the theoretical curves fitted to them. For this specimen $f_m = 1{\cdot}12$ kc/s and $a = 0{\cdot}0519$ cm, so that from (4), $\rho = 0{\cdot}0378\ \mu\Omega$ cm.

Discussion

If the resistivity of the specimen is not isotropic, the value of ρ measured by the present method is some average over directions normal to the specimen axis. If the specimen is a single uniaxial crystal, with principal resistivities $\rho_{\|}$ and $\rho_{\perp}$ parallel and perpendicular to the crystal axis, and the anisotropy is not too great, the apparent resistivity is approximately (Fraser and Shoenberg 1949)

$$\bar{\rho} = \tfrac{1}{2}\{\rho_{\perp} + (\rho_{\perp}\cos^2\phi + \rho_{\|}\sin^2\phi)\},$$

where ϕ is the angle between the specimen axis and the crystal axis. Alternatively, if a specimen were used in the form of a long thin strip, with one of the principal axes of the material (the k axis say) parallel to the width of the strip, the apparent resistivity would be approximately ρ_k. The details of the above analysis, of course, would have to be replaced by that appropriate to the altered geometry.

The range of resistivities that can be measured by the present method depends on the frequency range available and the bridge sensitivity. With simple apparatus, measurements are not readily extended below about 20 c/s, so that if anomalous skin-effect difficulties are to be avoided the lowest resistivity that can be measured is about $0{\cdot}002\ \mu\Omega$ cm, from (6). To avoid size-effect errors, f_m should then be less than 1 c/s, from (7); i.e. a should be greater than 0·4 cm. Under these conditions, however, the measured resistivity will be that of the surface layers rather than the bulk material, since $f \gg f_m$. Size effect errors will probably not be serious if a specimen of radius 0·1 cm is used; then $l/a \sim 0{\cdot}05$ and $f_m \sim 20$ c/s so that $f \sim f_m$ and the measured resistivity will be that of the bulk material. Taking a coil constant M_c/A_c of 400 μH cm^{-2}, a specimen of resistivity $0{\cdot}002\ \mu\Omega$ cm will then give an impedance change of about 0·5 mΩ from (5); large enough to measure with fair accuracy.

For materials of very high resistivity, such that $f \ll f_m$ even at the highest measuring frequency, we can use the approximations $m' = x^4/48$, $m'' = x^2/8$, and we then find for the impedance change on inserting the specimen:

$$\omega\delta M' = (M_c/A_c)8\pi^6 a^6 f^3/3\rho^2,\ \omega\delta M'' = (M_c/A_c)2\pi^4 a^4 f^2/\rho.$$

Taking $f = 20$ kc/s, $M_c/A_c = 400\ \mu\text{H cm}^{-2}$, and $a = 0{\cdot}1$ cm again, we find that $\omega\delta M' = 1$ mΩ for $\rho \sim 10^2\ \mu\Omega$ cm, and $\omega\delta M'' = 1$ mΩ for $\rho \sim 3 \times 10^3\ \mu\Omega$ cm. Thus if an impedance change of 1 mΩ is measurable at 20 kc/s, method (*c*) can be used for resistivities up to $10^2\ \mu\Omega$ cm, and method (*a*) up to $3 \times 10^3\ \mu\Omega$ cm. These limits can be greatly extended by increasing the specimen radius above 0·1 cm, though very high resistivities can be more simply measured by dielectric loss measurements.

A closely related method of measuring resistivity has recently been described by Bean, de Blois and Nesbitt (1959), in which the rate at which magnetic flux escapes from a conducting cylinder is measured, after an external steady field has been switched off. The rate of flux expulsion

determines the e.m.f. induced in a pick-up coil wound round the specimen. After a time of the order of τ, the induced e.m.f. decays exponentially with a time constant $\tau = 1 \cdot 10/2\pi f_m$. Using suitable equipment, values of τ between 50 s and about $0 \cdot 5$ μs can be measured, corresponding in f_m to $0 \cdot 003$ c/s and 300 kc/s, and for a specimen 1 cm in diameter this corresponds to a range of resistivities from 10^{-5} $\mu\Omega$ cm to 10^3 $\mu\Omega$ cm. This method is therefore particularly powerful for materials of very low resistivity, though not quite so powerful as this example suggests: a sample of resistivity 10^{-5} $\mu\Omega$ cm would have $l \sim 1$ cm, so that size-effect errors would be considerable in a rod 1 cm in diameter. Moreover, the radius of the electron orbits in typical metals (in cm) is related to the field H (in oersted) by $r \sim 5/H$ so that further complications due to orbit curvature and magnetoresistance would arise in fields of 1 oersted or more. A safer lower limit for this method would be 10^{-4} $\mu\Omega$ cm. (Since the decaying e.m.f. contains appreciable Fourier components up to at least $10 f_m$, condition (6) for avoiding anomalous skin-effect errors becomes identical with condition (7) for avoiding size-effect errors: we need at least $l/a < 10^{-1}$ and preferably $l/a < 10^{-2}$, to avoid such errors.) It remains true that for materials of the lowest resistivity the flux decay method is superior, but for resistivities above 10^{-2} $\mu\Omega$ cm, the mutual inductance method as discussed in the present paper may be simpler and more convenient.

References

BEAN, C. P., DE BLOIS, R. W., and NESBITT, L. B., 1959, *J. Appl. Phys.*, **30**, 1976.

CHAMBERS, R. G., 1952, *Proc. Roy. Soc.* A, **215**, 481.

CHAMBERS, R. G., and PARK, J. G., 1960, *Proc. VII Int. Conf. Low Temp. Physics, Toronto.*

FRASER, A. R., and SHOENBERG, D., 1949, *Proc. Camb. Phil. Soc.*, **45**, 680.

GRUBE, G., and SPEIDEL, H., 1940, *Z. Electrochem.*, **46**, 233.

KOUWENHOVEN, W. B., and DAIGER, G. P., 1934, *Rev. Sci. Instrum.*, **5**, 94.

LAURMANN, E., and SHOENBERG, D., 1949, *Proc. Roy. Soc.* A, **198**, 560.

REUTER, G. E. H., and SONDHEIMER, E. H., 1948, *Proc. Roy. Soc.* A, **195**, 336

RORSCHACH, H. E., and HERLIN, M. A., 1951, *Phys. Rev.*, **81**, 467.

SHOENBERG, D., 1952, *Superconductivity* (Cambridge University Press), p. 237.

VAN DEN BERG, G. J., and VAN DER MAREL, L. C., 1954, *Proc. I.I.R., Grenoble*, p. 161.

Notes and Comments

The Second International Conference on Stress Analysis

The Second International Conference on Stress Analysis organized by the Groupment pour l'avancement des Methods d'Analyse des Contraintes (G.A.M.A.C.) will be held in Paris from 10th–14th April 1962. It is not intended to include review papers in the Conference. Contributions will deal mainly with the results of experimental work, and there will be no publication of the complete texts by G.A.M.A.C. The Joint British Committee for Stress Analysis is arranging for the scrutiny of the British contributions on behalf of G.A.M.A.C. Abstracts and summaries must be submitted by 30th September 1961. Further information may be obtained from the Secretary, The Joint British Committee for Stress Analysis, The Institution of Mechanical Engineers, 1 Birdcage Walk, Westminster, S.W.1.

Symposium on Carbohydrate Chemistry

An International Symposium on Carbohydrate Chemistry, sponsored by the Chemical Society, will be held in Birmingham from 10th–20th July 1962. Further particulars will be published by the Society in due course. Copies of the announcement may be obtained, when available, from the General Secretary, The Chemical Society, Burlington House, London, W.1. Papers read at this meeting will not be published in collected form.

Third International Symposium on Rarefied Gas Dynamics

The Third International Symposium on Rarefied Gas Dynamics will be held at the University of Paris, on 26th–29th June 1962. The programme will range from topics of immediate significance for upper atmosphere and space flight to basic scientific studies and will include the following areas: Studies of the limits of the continuum theory or the quasi-equilibrium kinetic theory of gases; Problems in kinetic theory of gases, particularly attempts to solve the Boltzmann equation; Free-molecule and near-free-molecule flow in neutral and ionized gases; The physics of surface interactions between gases and solids; Boundary conditions for rarefied gas equations—slip flow; Experimental techniques and instrumentation developments bearing on the above, whether applied to laboratory or field experiments.

Inquiries from the U.S. should be addressed to L. Talbot, Department of Aeronautical Sciences, University of California, Berkeley, and from Europe to Laboratoire d'Aerothermique, 4c, route des Gardes, Meudon (S. and O.), France.

Photoconduction in cadmium sulphide

by J. F. DUNCAN, M.Sc., M.A., D.Phil., and D. N. SITHARAMA RAO, M.Sc., Ph.D., University of Melbourne, Chemistry Department, Carlton, N.3, Victoria, Australia

MS. received 27*th January* 1961, *in revised form* 15*th March* 1961

Abstract

Discrete voltage pulses, of amplitude depending on the energy of the incident radiation have been observed in the photoconduction of cadmium sulphide at room temperature. Between 4800 Å *and* 8000 Å *the pulse height varies in a similar manner to the photocurrent with wavelength. Below* 3500 Å *it is closely proportional to the energy of the radiation. Cadmium sulphide can be used as a spectrophotometric device in this region when only small numbers of quanta are available.*

PHOTOCONDUCTION in heavy metal sulphides, tellurides, and antimonides, and in silver halides has been known since 1920 (Gudden and Pohl 1920, Breckenridge *et al.* 1956, Gibson 1958), and the d.c. response is currently usefully employed for detection of low intensities of radiation. Likewise, a liquid-air-cooled single crystal can be used for detection of nuclear radiation by counting the voltage pulses produced when nuclear particles are allowed to impinge on its surface (Hofstadter 1949). We have been interested in these detectors for investigating the low energy emission obtained during external gamma irradiation of suitable substances, and during nuclear decay. But the few crystal spectrometers (Hofstadter 1949) which have been previously reported have been sensitive only in the high energy region. We now report some of the results obtained during our search for a device which could respond effectively to single quanta in the region between 1 and 100 ev.

Cadmium sulphide was prepared by the method of Frerichs (1947) from technical grade cadmium and hydrogen sulphide. Emission spectrographic analysis of the resulting crystals showed no detectable impurity which could not be ascribed to the known impurity content of the carbon specimen rods. In particular bismuth, copper, gallium, indium, iron, and magnesium were present in less than 0·01 parts per million. Analysis for cadmium and sulphur by the formation of the oxides revealed that the crystals were almost stoichiometric. There was, however, some evidence that the cadmium content varied, just outside the limits of experimental error (see Table 1).

Table 1. *Ratio of cadmium to sulphur in cadmium sulphide crystals*

Temp. of preparation (° C)	*Batch* 1	*Batch* 2	*Batch* 3
800	1·00	1·00	1·00
900	1·04	1·04	1·06
1000	1·03	1·00	1·00

The experimental accuracy in these figures is ±0·02.

The crystals were mounted either in platinum spring contacts, or with contacts made by soldering the electrical leads with indium metal. The latter gave good results, whereas the former sometimes exhibited spurious pulses arising from the contact. The resistances of the crystals are listed in Table 2, together with the changes observed when irradiated with about 10^{15} quanta/sec of white light. A small excess of free cadmium in the lattice of the cadmium sulphide would be expected to have a profound effect on the properties of the crystals, since it would increase the number of free charge carriers. Crystals prepared at the same temperature, and expected to have similar composition, had very similar resistances. But those prepared at 900° C usually had a slightly higher dark conductivity, and the change in d.c. on illumination was much smaller than in those prepared at higher or lower temperatures. They were, however, more suitable for investigation of the pulse amplitude response using monochromatic visible radiation since the pulses were less frequent, and more easily observed. Nevertheless, although the response of these crystals to individual quanta was less efficient than those prepared at other temperatures, the observed counting rate increased linearly with intensity up to about 60000 counts/min. The pulse amplitude also increased linearly with increasing voltage up to a field strength of 6–800 v/cm, but thereafter became less dependent on voltage, until above 1200 v the pulse amplitude was independent of voltage. The results reported here were determined using an operating voltage of 500 v/cm. Hall coefficient measurements on all crystal preparations confirmed that they were n-type semiconductors.

Table 2. *Electrical response of cadmium sulphide crystals*

Temp. of preparation (° C)	*Resistance* (Ω cm) *Typical crystals* *Dark*	*Photo*	*Range for all crystals* *Dark*	*Photo*
800	$4·1 \times 10^7$	$1·3 \times 10^4$	10^7–10^9	10^2–10^4
	$6·4 \times 10^8$	$4·3 \times 10^2$		
	$5·9 \times 10^8$	$4·1 \times 10^3$		
900	$2·4 \times 10^6$	$1·3 \times 10^6$	10^5–10^8	10^3–10^6
	$5·7 \times 10^7$	$3·2 \times 10^7$		
	$5·0 \times 10^5$	$3·2 \times 10^3$		
1000	$8·2 \times 10^7$	$1·1 \times 10^6$	10^7–10^8	10^4–10^6
	$5·5 \times 10^7$	$1·8 \times 10^6$		
	$1·6 \times 10^7$	$8·2 \times 10^4$		

Crystals which showed good photoconduction were illuminated with monochromatic radiation from an Optica spectrophotometer. Two quantities were measured as a function of wavelength, namely the total current, and the pulse height produced by individual quanta when amplified by an Ekco linear amplifier type 1049B and displayed on a cathode-ray oscilloscope (see Fig. 1). Not all the crystals exhibited good pulses, or gave pulses of uniform height. But in favourable cases, all pulses were of the same amplitude for the same wavelength, and the response was such that the

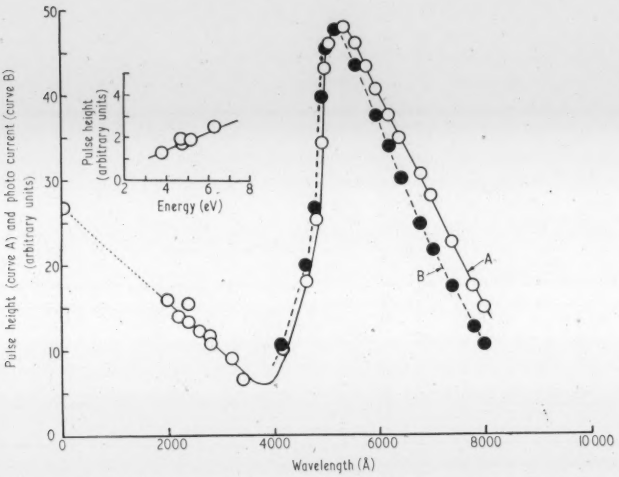
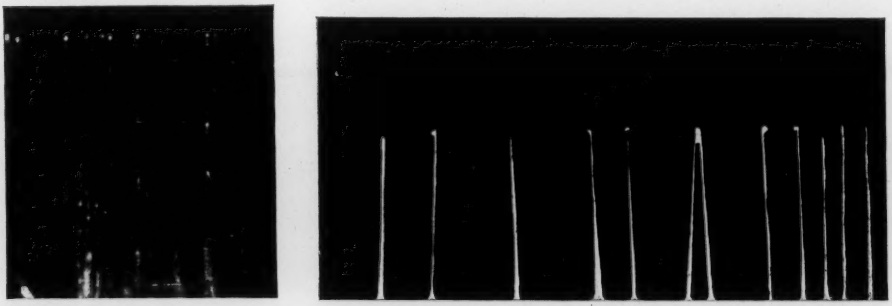


Fig. 1. Pulse height, curve A, and photocurrent, curve B, of a typical cadmium sulphide crystal from the preparation at 900° C, as a function of the wavelength of monochromatic radiation. Between 1850 Å and 8000 Å the Optica spectrophotometer was used. Points were also obtained at 1·93 Å from a radioactive ^{55}Fe source, and at 0·75 Å from zirconium K x-radiation produced by activation of zirconium foil with a beta active (^{14}C) radioactive source. The inset shows the region between 1800 Å and 3500 Å plotted as a function of energy.

crystal could be used as a spectrophotometric device. Fig. 2 illustrates some of the pulses recorded.

Figure 1 displays three regions in which photoconduction may result. At the peak (5200 Å) the quantum energy corresponds to *one* photon of 2·4 ev which would be only just sufficient to raise *one* electron from the filled band into the conduction band in a stoichiometric cadmium sulphide crystal (Kröger *et al.* 1954). The fact that single pulses of large and variable amplitude are obtained in this region between 5200 Å and 8000 Å shows this crystal to contain many low-lying impurity levels. Likewise, a voltage pulse of 3×10^{-4} v obtained at 5200 Å shows that multiplication occurs in the crystal. This is not uncommon. Bube (1955), for instance, reports a gain of 10^4 in some cadmium sulphide crystals on irradiation.

In the second region between 4000 Å and 5200 Å there is a monotonous increase in pulse height with wavelength. This can be correlated with the very large absorption



(*a*) (*b*)

Fig. 2. Cathode-ray oscilloscope display of pulses from the crystal used to obtain Fig. 1.

(*a*) 4800 Å (2·58 ev). This is an untouched photograph of the cathode-ray oscilloscope.

(*b*) ^{55}Fe radioactive source (6·4 kev). In this case, the pulses were not of sufficient intensity to produce a good reproduction in the journal. The background has therefore been blackened, and the pulses painted with white ink to produce good contrast, before photographing a second time.

In both photographs, the full screen sweep corresponds to 10 μs. Only half of the screen is shown in Fig. 2(*a*).

coefficient of cadmium sulphide below 5200 Å which requires that most of the radiation loss occurs near the surface (Frerichs 1947). By contrast, however, in the third region, below 3600 Å, the penetration of the crystal by the radiation becomes increasingly large as the energy increases. The number of primary electrons formed must also increase as the wavelength is decreased, so that at the higher energies larger numbers of secondary electrons are produced by collision with neutral cadmium atoms as the primary electrons are accelerated under the applied voltage. The role of secondary electron production is clearly shown (Kallman and Warminsky 1948) by the very large voltage pulses produced from a particle of high energy and low penetration power (such as an alpha particle). Fig. 1 shows that the pulse height for electromagnetic radiation is linear with energy at least from 3·8 to 8·0 ev. Whether it is strictly proportional to energy over an extended range depends on the operating voltage and the dimensions of the crystal, just as it does with a gas-filled proportional counter. If strictly linear, the plot of Fig. 1 would extrapolate to zero pulse height at an energy in the region of 1 ev. This confirms the conclusion of the previous paragraph that the crystal contains many low-lying impurity levels.

Our results demonstrate that (*a*) the electrical response of cadmium sulphide depends on the energy of the radiation and the method of preparation, (*b*) the response to single quanta by cadmium sulphide crystals in the region of 5200 Å involves considerable electron multiplication, and is dependent on the impurity level, the absorption characteristics and other features, and (*c*) suitably prepared crystals can be used as spectrophotometric devices at room temperature.

Acknowledgments

One of us (D. N. S.) thanks the Anti-Cancer Council of Victoria for a postdoctoral research fellowship, tenable with the Radiochemical Group of the University of Melbourne.

References

BRECKENRIDGE, R. G., HAHN, E. H., and RUSSELL, B. R. (Ed.), 1956, *Photoconductivity Conference* (New York: Wiley).
BUBE, R. H., 1955, *J. Chem. Phys.*, **23**, 15, 18.
FRERICHS, R., 1947, *Phys. Rev.*, **72**, 594.
GIBSON, A. F. (Ed.), 1958, *Progress in Semiconductors*, Vol. 3 (Manchester: Heywood).
GUDDEN, B., and POHL, R., 1920, *Zeit. fur Physik*, **2**, 181.
HOFSTADTER, R., 1949, *Nucleonics*, **4**, No. 4, 2.
KALLMANN, H., and WARMINSKY, R., 1948, *Ann. Phys.*, **4**, 69.
KRÖGER, F. A., VINK, H. J., and VOLGER, J., 1954, *Physica*, **20**, 1095.

Notes and comments

New publications of the National Bureau of Standards

We have received notice of the following new publications of the National Bureau of Standards: *Atomic Energy Levels in Crystals* by John L. Prather, National Bureau of Standards Monograph 19, 84 pages, price 60 cents. This monograph presents a theoretical study of the sharp line absorption spectra of crystals. *Corrected Optical Pyrometer Readings* by E. E. Poland, J. W. Green and J. L. Margrave, National Bureau of Standards Monograph 30, 74 pages, price 55 cents. This table can be used to convert the observed temperature to the true temperature, taking into account the effective emissivity of the material, or conversely to determine the effective emissivity if the true temperature is known. *Ideal Gas Thermodynamic Functions and Isotope Exchange Functions for the Diatomic Hydrides, Deuterides and Tritides* by Lester Haar, Abraham S. Friedman and Charles W. Beckett, National Bureau of Standards Monograph 20, 271 pages, price $2.75.

These monographs may be obtained from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C.

Canadian Nuclear Technology

We have received notice of a new Canadian publication, *Canadian Nuclear Technology*, which is published quarterly starting from June this year. It will report on significant scientific achievement in Canada and throughout the world. An editorial advisory board, consisting of senior professional men in the nuclear industry, will referee the technical editorial content of the journal. Further information may be obtained from MacLean-Hunter of Canada, 30 Old Burlington Street, London, W.1.

New British Standards

The following new British Standards have recently been issued: B.S. 3383: 1961. Normal equal-loudness contours for pure tones and normal threshold of hearing under free-field listening conditions. Price 4*s*. 6*d*. B.S. 1828: 1961. Reference tables for copper v. constantan thermocouples. Price 6*s*. These may be obtained from the British Standards Institution, British Standards House, 2 Park Street, London, W.1.

Rheology of granular material II—a method for the determination of the intergranular cohesion

by M. M. BENARIE, M.Sc., DSc., Ing.Chem., F.P.S., Israel Institute for Biological Research, Ness-Ziona, Israel

MS. received 7th February 1961, *in revised form 24th April* 1961

Abstract

The intergranular cohesion of granular masses is investigated with an apparatus which in principle is a Couette viscometer. Rheologically, the granular mass behaves as a plastic solid, with the difference that the plastic yield point S_0 of a granular mass is a function of the pressure p which acts normally on the yielding plane. In an (S_0, p) diagram, a cohesionless mass will be represented by a straight line passing through the origin and having a limiting slope of 0·193, while a granular mass with cohesion will show a lesser slope, the departure from the limiting value being a convenient measure of the cohesion. Powders which are not free-flowing will be represented in the (S_0, p) diagram by lines, which are not necessarily straight, with an intercept on the S_0 axis.

The cohesion of a few samples of sand of different granularity and that of a few other granular masses has been measured.

1. Introduction

While the *adhesion* of powders and granular masses to various surfaces has been the subject of a few investigations (Buzagh 1929, Derjagin and Lasarew 1935, Patat and Schmid 1960), the other, technically not less important, aspect, the *cohesion* has not been studied systematically. The 'free-flow', 'dustability', 'dispersability' or 'stickiness' of a powder are intuitive and technical expressions for its behaviour under the influence of cohesive and cohesion-like forces; they very rarely lend themselves to quantitative description.

A granular mass is a rather complicated system which is defined only by a complete description of the physical nature of the grains which form the mass (their strength, elasticity, etc.), the form of the grains, the state of the grain surfaces, to mention only a few of the most important factors without attempting a comprehensive description. The present paper is limited to the examination of the cohesion-like intergranular forces on an idealized granular mass. Its properties are defined as follows: the mass is a conglomerate of perfectly rigid and—at least under the conditions of the experiment—unbreakable particles in contact with each other. Homogeneous dispersion and a grain form which excludes interpenetration, mechanical accretion or occlusion are further essential features of the idealized mass.

In the absence of cohesion, the ideal granular mass behaves rheologically as a plastic body (Endersby 1940), but is different from an ideally plastic solid (Bingham body) in that the plastic yield point of a granular mass is a function of the pressure which acts normally on the yielding plane (Benarie 1960). The main purpose of the present work is to confirm experimentally and to explain theoretically the proportionality between plastic yield point and normal pressure, and to establish the limits within which this proportionality may be considered valid.

In studying the work required to displace a layer of granular mass relative to another layer of the same mass, it is convenient to separate the factors which are already known to be proportional to the pressure from those which are either not fully known or for which the above proportionality has not yet been demonstrated.

The work against the 'classical' Coulomb friction (solid-to-solid friction) is by definition proportional to the pressure; the proportionality also exists for the work expended to dilate the granular mass (Reynolds 1885) so as to allow deformation or displacement. Any experimentally observed deviation from proportionality must, therefore, be ascribed to some other factor or factors. Cohesion which obviously interferes with the relative displacement of the grains of a mass may be present in one or more of the following forms:

(*a*) Structural cohesion, such as, for example, the formation of chains or other interconnected forms from grains possessing hook-like protrusions. This factor, as stated above, is not included in the definition of the ideal granular mass.

(*b*) Homogeneous bonding, in form of the London–van der Waals forces (Bradley 1932). This is the cohesion resulting from the surface energy acting among solid particles of the same material.

(*c*) Heterogeneous bonding, ascribed to the adsorbed or otherwise bound layer of gas or liquid on the grains. This effect is practically always present. Intergranular layers of liquid will not only affect cohesion, they even may transform the solid-to-solid friction (which is independent of the velocity gradient) into viscous friction, which is a function of the velocity gradient.

It is possible that other lesser known or less important forces also counteract the displacement of one layer of a granular mass relative to the neighbouring layer. In the following discussion, therefore, the expression 'cohesion-like force' will be used instead of 'cohesion' for the whole group of factors which counteract the deformation of a granular mass and are *not* directly proportional to the pressure.

2. Apparatus

The apparatus used was, in principle, a Couette viscometer. The plastic yield point of the granular mass was determined by rotating an appropriately surfaced cylinder immersed in the granular mass and measuring the torque needed for the rotation. The cylinder surface must permit the formation of a boundary layer adherent to the cylinder, thus permitting the measurement of the intergranular friction in the mass.

Instead of a rotating cylinder, a plane could be used provided that the plane retains an immobile boundary layer;

it is slowly inclined until a part of the granular layer begins to slide. However, this experimental array, used for the determination of adhesion of powders to solid surfaces by Cremer *et al.* (1952) and Patat and Schmid (1960), has not been adopted in the present study because (*a*) we attempted to determine the cohesion under controlled conditions of relative movement and not only at the discontinuity as the movement begins, (*b*) the packing condition of the powder poured on to the plane is not well defined nor easily measured, and (*c*) the method of the inclined plane does not lend itself to experimentation under widely variable pressures.

Three forms of rotors were used for the measurement of the torque caused by friction, the section through two of them being shown in Fig. 1. The third rotor was a simple brass cylinder, coated with an adherent layer of emery cloth of the desired coarseness.

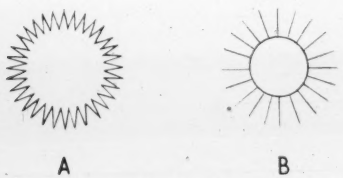


Fig. 1. Section through the rotors.

The form of the functional dependence of the torque on granularity, pressure and other variables is independent of the type of rotor, provided that an adherent layer is present on the rotor. The absolute values of the torque are different when measured with rotors of different type, but equal diameter, because of the different thickness of the adhering layer. The diameter of the rotor plus twice the thickness of the adhering layer is called the effective diameter d_{eff}; it can easily be measured by means of callipers, the rotor being half immersed. The accuracy obtained is to $\pm 0{\cdot}02$ cm.

Table 1 presents a comparison between the real and effective diameters of the rotors as measured in sand of 45–50 mesh. The last line shows the effective diameters as calculated from the mean of all torque measurements made during the present work in sand of 45–50 mesh. For the calculation, the effective diameter of rotor No. 1 was chosen as an arbitrary unit.

Table 1. *Comparison between the real and effective diameters of the rotors in use*

	Rotor No. 1 A (Fig. 1)	*Rotor No. 2 B (Fig. 1)*	*Rotor No. 3*
Real external diameter	32·0 mm	32·0 mm	33·2 mm
Measured effective diameter	34·0 mm	38·6 mm	40·6 mm
Effective diameter, calculated from torque measurements taking the effective diameter of rotor No. 1 as unity	1·00	1·15	1·20

The rotor revolves in the granular mass; angular velocities from 1 rev/min to 10 rev/min may be applied, using a gear-box between the motor and the rotor. The torque is measured by the torsion of a piece of latex tubing.

By frequent recalibration with known weights the precision and reproducibility of this rather primitive device was maintained better by a factor of at least five than would usually have been the case with such materials which are so inhomogeneous and whose packing is so ill-defined. The error of the calibration was less than ± 10 g cm, while the deviation of any single torque measurement from the mean obtained in the same conditions was often ± 100 g cm.

The torque values used for the computation of the cohesion-like forces in the granular mass are always the equilibrium torques, T_∞, measured after attaining constant and reproducible packing density in the material. If the granular material is poured slowly into the container and over the rotor, the starting torque has a low initial value and rises exponentially until T_∞ is attained. As containers, standard Pyrex beakers of 3 l. (about 15 cm diameter and 20 cm high) were used. It is thought that beakers of these dimensions practically eliminate any wall effect of the container.

If the beaker containing the granular material, with the rotor in place, is shaken well (e.g. on a sieve shaker) before the torque measurement, thus obtaining the greatest practically attainable packing density ($\rho = 1{\cdot}68$ g/cm^3 in case of the quartz sand used, as against $\rho = 1{\cdot}50$ g/cm^3 in the case of the same sand in 'poured' condition), the starting torque has high values and diminishes exponentially until, with the same material, the same T_∞ is attained as previously. This behaviour is graphically illustrated by Fig. 2.

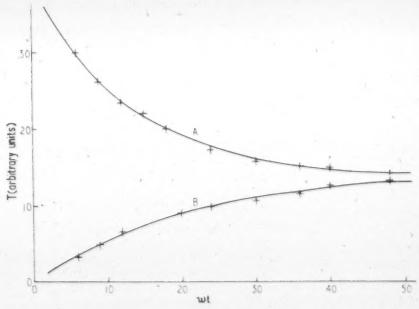


Fig. 2. An example of the asymptotic approach to the equilibrium torque T_∞. Torque against time representation. Curve A, $(T - 14)/26 = \exp(-0{\cdot}081\ \omega t)$; curve B, $(14 - T)/14 = \exp(-0{\cdot}051\omega t)$; $\omega = 3$ rev/min, t in minutes. Sand 50–70 mesh. + Experimental points.

An empirical dependence of the torque on the duration of the experiment, with low density (dilated) as well as with high density (settled) material can be satisfactorily represented by the relation:

$$|T - T_\infty| / |T_\infty - T_0| = \exp(-k\omega t)$$

where T is the torque at any time, T_0 the torque at the beginning of the experiment and T_∞ the equilibrium torque attained after a very long time, k is an apparatus constant, depending on the diameter and surface condition of the rotor, ω the angular velocity and t the duration of the experiment. This relationship was satisfactorily verified with every granular mass used and for any angular velocity between 1 and 10 rev/min. In a limited measure, the relationship can also be used to find T_∞ by extrapolation. This was especially the case with the lowest angular velocities, as at 1 rev/min, T_∞

was attained only after twenty or more hours of rotation. In these cases, T_∞ was extrapolated from the readings made during the first few hours of experiment, and then the apparatus left unattended for several more hours, when confirmatory readings were taken.

The assumption that the measured equilibrium torques are characteristic of the density is substantiated by the fact that the rotation may be suspended for several hours (longer rest periods than 96 hours were not tried), either before or after arriving at the asymptotical value of torque without any change in the torque value before or after the stoppage. The curve representing the variation of T with t can be traced without any discontinuity, when t represents the duration of the actual rotation of the rotor. The interruptions, even when more than one, may be disregarded. Even interruptions lasting a multiple of t are without any influence on the result.

3. Theory

The theoretical question is how to evaluate the pressure which acts normally on the plane of failure in the granular mass. Since in the Couette-type apparatus the yielding plane is the mantle of a vertical cylinder, we are interested mainly in the horizontal pressure. According to Rankine (1856), the horizontal pressure p_h is proportional to the vertical pressure p_v as shown by the relation:

$$p_h = p_v \frac{1 - \sin\phi}{1 + \sin\phi} = Rp_v$$

where ϕ is the angle of repose of the material. The experimental verification of the proportionality shown by Rankine's relation was one of the main objects of the present paper.

The fact that the torque was independent of the angular velocity was experimentally ascertained in a large number of experiments; thus the flow line at any given pressure is a straight line parallel to the shear rate axis. The magnitude of the intercept on the shear stress axis, in the case of a non-cohesive mass, is proportional to the pressure. From this form of shear rate–shear stress curve it follows directly that the flow of sand through capillary tubes most be largely independent of the head, as was found by Bingham and Wikoff (1931).

From the theory of the rotating cylinder instrument it follows (Reiner 1949) that, in general, the torque–angular velocity diagram must have the same form as the shear rate–shear stress diagram. As in the case of an ideal granular mass the torque is independent of angular velocity, the graph will be represented by a line parallel to the angular velocity axis, the intercept on the torque axis being $2\pi d_{\text{eff}}^2 L S_0$, where L is the height of the cylinder and S_0 the plastic yield stress.

The torque caused by friction on the rotating cylinder is given by the effective surface of the cylinder $\pi d_{\text{eff}} L$, multiplied by the mean pressure acting on the cylinder $R\mu\rho(\frac{1}{2}L + h)$, where R is the supposed proportionality factor between vertical and horizontal pressures, μ the coefficient of self-friction of the granular mass ρ its specific gravity and h the height of the mass above the upper rim of the rotating cylinder, and further multiplied by the effective radius $\frac{1}{2}d_{\text{eff}}$.

$$T = \tfrac{1}{2}\pi\mu R\rho d_{\text{eff}}^2 L(\tfrac{1}{2}L + h) \qquad (1a)$$

or

$$S_0 = \frac{2T}{\pi d_{\text{eff}}^2 L} = \mu[R\rho(\tfrac{1}{2}L + h)]. \qquad (1b)$$

The torque and the plastic yield stress are then proportional to the mean pressure, the latter being the expression appearing in the square brackets.

μ the friction coefficient of the granular mass is given by $\mu = \tan\phi$. Let us suppose that

$$mg \sin\phi = \mu_1 (mg \cos\phi + C) \qquad (2)$$

where m is the mean mass of a grain, g the acceleration of gravity, μ_1 and C constants, the latter accounting for the intergranular cohesion-like forces. This equation was used by Derjagin and Lasarew (1935) to express the *adhesion* of mineral powders to different surfaces.

From Eqn (2),

$$\mu_1 = \frac{mg \sin\phi}{mg \cos\phi + C} = \mu \frac{mg \cos\phi}{mg \cos\phi + C}.$$

In the case of positive attraction among the grains—cohesion-like forces—C is positive and $\mu > \mu_1$. If the cohesion-like forces are tending to zero, $C \ll mg \cos\phi$ and $\mu_1 \simeq \mu$.

If cohesion-like forces are present μ should be replaced by μ_1 in Eqn (1):

$$S_0 = \frac{2T}{\pi d_{\text{eff}}^2 L} = \mu \frac{mg \cos\phi}{mg \cos\phi + C} R\rho(\tfrac{1}{2}L + h) \qquad (1c)$$

and the slope of the line representing S_0 against the pressure is smaller than in the absence of cohesion-like forces.

$$\frac{2T}{\pi d_{\text{eff}}^2 \rho L(\frac{1}{2}L + h)} = \frac{\sin\phi}{\cos\phi} \cdot \frac{(1 - \sin\phi)}{(1 + \sin\phi)}. \qquad (3)$$

If for a granular mass the validity of Eqn (3) (which is only the explicit form of Eqn (1*b*)) is verified, then the mass is ideal and cohesion-like forces are absent. If experimental evidence shows that the right hand side of Eqn (3) has to be multiplied by a factor, $f = mg \cos\phi/(mg \cos\phi + C)$ to account for the facts, an inter-grain cohesion-like force is present, which amounts to

$$C = mg \cos\phi \frac{1 - f}{f} \qquad (4)$$

for a grain of mass m.

When a monodispersed granular mass is behaving ideally, one would expect that following the definition given in the Introduction, every grain has the same mean number of contact points with other grains (as, for example, spheres have in a given regular arrangement), and each contact point contributes equally to the total cohesion. By this rough idealization (arrived at mainly by neglecting the fact that cohesion will depend also on the curvature of the surfaces in contact) one is led to the conclusion that the cohesion of each grain is independent of the grain dimensions. If these assumptions were strictly true, the factor $(1 - f)/f$ would be independent of the grain dimensions. Actually the cohesion-like force as expressed by C is a characteristic of the material and depends on the form, distribution and dimension of the grains and has to be experimentally measured for each material.

Given that in the present determination of the cohesion-like force, the stress of the granular mass is mainly due to shear—although some grain translation and rotation also takes place—the constant C is a characteristic for the shearing behaviour of the mass.

Following Eqn (1*c*), the idealized behaviour of the granular mass can be schematically represented in an (S_0, p) diagram, as shown in Fig. 3.

The greatest possible numerical value of tan α (the slope corresponding to $S_0 = \alpha p_v$) is 0·193, because

$$\alpha = f \tan \phi \, (1 - \sin \phi)/(1 + \sin \phi)$$

and in the absence of cohesion f is equal to unity; the function representing α has a maximum at $\phi = 30°$.

If intergranular cohesion is present, the behaviour of the mass will be represented by a line B in Fig. 3, with $\alpha' < \alpha$.

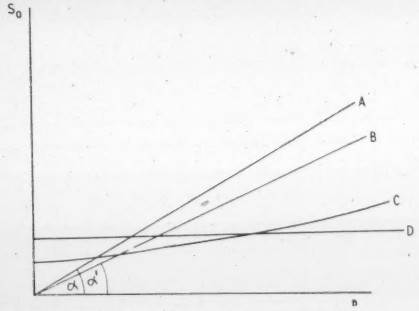


Fig. 3. Schematic representation of the yield stress S_0 against pressure p: A, cohesionless mass; B, cohesive but free-flowing mass; C, sticky, i.e. not free-flowing powder; D, Bingham body.

In the course of the present investigation it was repeatedly established that granular masses which possess an intercept on the S_0 axis cease to be free-flowing. Inversely, the presence or absence of an intercept on the S_0 axis may be considered as a quantitative measure of the 'free-flow' of a granular mass.

The representation of a Bingham body in Fig. 3 would be the straight line D, which is parallel to the pressure axis. This relation explains why it is insufficient to consider a granular mass simply as a plastic body. While a Bingham body is represented by the usual rheological model of friction element, this is insufficient for the granular mass, which can only be represented by a model containing a new element, the pressure dependent friction (Benarie 1960).

4. Results

The main difficulty in the course of the measurements was caused by the settling and the dilatation of the material. Settling causes the formation of a depression, while dilatation causes the appearance of a ridge over the rotor rim. The depression has to be carefully filled in by pouring some more material into it; the ridge has to be smoothed, before significant pressure head (height) measurement can be made over the rotor. The determination of the height of the granular mass over the rotor was made repeatedly (5 to 10 times) for each T_∞ value, and the mean was taken. The end correction for the rotor was empirically established by rotating a thin (0·2 mm) rotor of the same material and finish as the rotors used for the measurements. This correction did not amount in any case to more than 2% of the measured torque.

Figure 4 is a typical representation of the results, showing the dependence of the torque on pressure and grain size for a dry sand (mainly quartz) of well rounded grain, measured in air. One also obtains straight lines, but with other slopes, when the same sand is measured *in vacuo* (50 μ Hg pressure, over phosphorus pentoxide).

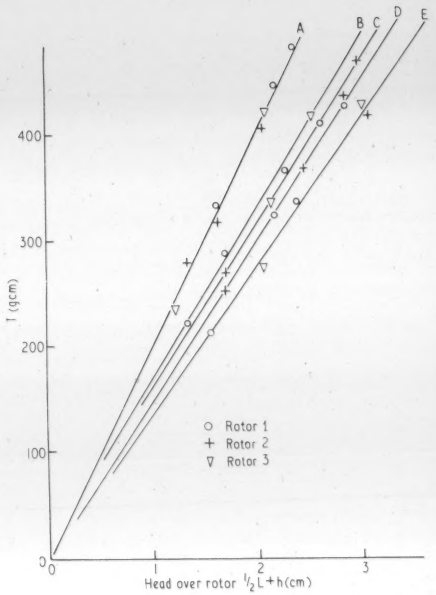


Fig. 4. Determination of the yield stress of dry sand of different granularity in function of the pressure; mesh: A, 45–50; B, 50–70; C, 100–120; D, 70–100, E, 120–140.

The same sand was coated with paraffin oil by rinsing the sand with a 0·1% solution of S.A.E. 30 oil in petrol and air-drying. Table 2 shows, that while the cohesion of the material *in vacuo* is slightly larger, but has the same dependence on the grain size as the cohesion in air, the cohesion of the oil-coated material is significantly greater.

Table 2. *Cohesion factor* $\{(1-f)/f\}$ *of sand in function of grain size*

Mesh	45–50	50–70	70–100	100–120	120–140
Dry sand in air	0·16	0·43	0·54	0·47	0·67
Dry sand *in vacuo*	0·47	0·56	0·70	0·75	0·82
Sand, oil coated	1·50	1·85	2·30	2·30	2·70

If the oil film on the grains (as computed from the oil weight per surface of sand) is thicker than 100 Å (the thickness of the oil film being considered uniform over the grain surface) the behaviour of the mass follows line C in Fig. 3 and, at the same time, the sand ceases to flow from a funnel with 3 mm opening although, macroscopically, one would call the sand 'dry sand'. If the oil coating wets the material so as to give the macroscopic impression of a wet sand, the torque becomes velocity dependent and the behaviour of the mass more complicated as viscous friction occurs between the grains. The behaviour of systems with viscous friction will be the subject of further study.

Carborundum from 40 to 120 mesh was also measured and found to be practically devoid of cohesion ($f > 0·97$).

Acknowledgments

The author wishes to express thanks to Dr. M. Bitron, Mr. G. Volinez and Mr. S. Kantorowitz for their interest and valuable discussions.

References

BENARIE, M. M., 1960, *Bull. Res. Counc. Israel*, **8F**, 229.
BINGHAM, E. C., and WIKOFF, R. F., 1931, *J. Rheology*, **2**, 395.
BRADLEY, R. S., 1932, *Phil. Mag.*, **13**, 853.
BUZAGH, A. V., 1929, *Kolloidzschr.*, **47**, 370.
CREMER, E., CONRAD, F., and KRAUS, T., 1952, *Angew. Chem.*, **64**, 10.
DERJAGIN, E., and LASAREW, W., 1935, *Colloid J., Voronezh*, **1**, 295.
ENDERSBY, V. A., 1940, *Proc. Amer. Soc. Test. Mater.*, **40**, 1154.
PATAT, F., and SCHMID, W., 1960, *Chemie-Ing.-Techn.*, **32**, 8.
RANKINE, W. J., 1856, *Phil. Trans. Roy. Soc. Lond.*, **146**, 9.
REINER, M., 1949, *Deformation and Flow*, p. 38 (London: Lewis).
REYNOLDS, O., 1885, *Phil. Mag.*, **20**, 469.

Notes and Comments

Conference on Imperfections in Crystals

The X-ray Analysis Group of The Institute of Physics and The Physical Society will hold its Autumn Conference on Imperfections in Crystals on 17th–18th November 1961, at The Institution of Mechanical Engineers in London.

The three sessions of the Conference will be devoted to the nature of crystal imperfections and of associated diffraction effects, to experimental methods for studying imperfections, and to their influence on the physical properties of crystals. Included in the invited speakers will be Dr. P. B. Hirsch, F.Inst.P., of the University of Cambridge, and Professor F. C. Frank, O.B.E., F.Inst.P., F.R.S., of the University of Bristol.

Correspondence regarding the programme should be addressed to the Honorary Conference Secretary, Dr. G. S. Parry, A.Inst.P., Department of Chemical Engineering, Imperial College of Science and Technology, Prince Consort Road, London, S.W.7.

Enquiries regarding attendance should be addressed to the Administration Assistant, The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1.

Digest of Literature on Dielectrics

A comprehensive international digest of technical publications on the theory and applications of dielectrics and related subjects has been prepared and published annually by the *Digest* Committee of the Conference on Electrical Insulation of the U.S.A. National Research Council. The 1959 *Digest*, prepared by a group of 38 experts, covers all aspects of dielectrics and treats the subject from the point of view of the physicist, the chemist and the engineer.

The *Digest* contains summaries and evaluations of technical advances in various phases of the field which were described in periodicals and books published in 1959. It consists of twelve chapters, each with an extensive bibliography, on different aspects of dielectrics: Instrumentation and measurements; Tables of dielectric constants, dipole moments and dielectric relaxation times; Molecular and ionic interactions in dielectrics; Conduction phenomena in solid dielectrics; The breakdown of dielectrics; Ferroelectric and piezoelectric materials; Magnetic materials; Rubber and plastic insulation; Insulating films and fibrous materials; Insulating liquids and their applications; Solid inorganic insulation; Applications.

The *Digest* consists of 421 pages, including a bibliography of 2,200 references. The price is $8.00.

Soviet Astronomy—AJ

Soviet Astronomy—AJ is a translation beginning with the 1957 issues of the *Astronomicheskii Zhurnal* of the USSR Academy of Sciences. It appears bimonthly and is published by the American Institute of Physics in co-operation with the National Science Foundation. The main subjects covered are: astrophysics and radio astronomy; stellar astronomy; the solar system; satellites, interplanetary medium and geophysics; and instrumentation.

The journal is available from the American Institute of Physics, 335 East 45th Street, New York 17, N.Y. The price is $27.

Spectral response of antimony–caesium photocathodes

by J. B. BIRKS, Ph.D., D.Sc., F.Inst.P., and I. H. MUNRO, B.Sc., Grad.Inst.P., The Physical Laboratories, The University, Manchester 13

MS. received 26th January 1961

Abstract

Measurements have been made of the relative photoelectric quantum efficiency $\eta(\lambda)$ of $SbCs_3$ cathodes of photomultipliers with Pyrex and quartz windows, from $\lambda = 200$–$650\ m\mu$. $\eta(\lambda)$ depends on the window transmittance $W(\lambda)$, the cathode absorbance $A(\lambda, d)$, the absolute photoelectric quantum efficiency $Q_0(\lambda)$ and the photoelectron escape probability $f(\lambda, d)$. Comparison with other data has allowed the separation of these factors, yielding inter alia *the absorption spectrum of $SbCs_3$ down to $\lambda = 220\ m\mu$, and the absolute photoelectric threshold curve $Q_0(\lambda)$ plotted against λ. The optimum thickness for a thin cathode is $d = 200$ Å. Errors in other published data are noted, and methods of increasing $\eta(\lambda)$ and of improving scintillation detectors are discussed.*

1. Introduction

ANTIMONY–CAESIUM ($SbCs_3$) is widely used as the cathode material of photomultiplier tubes and other photoelectric devices. It has a high photoelectric efficiency and low thermionic emission (dark noise) and it is sensitive in the blue–green region of the spectrum.

Photomultiplier tubes with thin $SbCs_3$ cathodes deposited on glass windows are commonly used in scintillation detectors. The S 11 spectral response of this particular cathode arrangement has strongly influenced the choice and design of scintillator systems to date. For example, secondary solutes are introduced into liquid and plastic organic solution scintillators mainly to shift the solute emission spectrum from the near ultra-violet into the blue–green to match the S 11 response. In a similar manner gas scintillators, which emit in the ultra-violet, are used in conjunction with 'wavelength-shifters, such as crystalline *p*-quaterphenyl, which absorb the primary emission and re-emit it in the blue–green. The choice of thallium as the impurity activator in the alkali halide crystal scintillators appears to be based mainly on the fact that its emission spectrum matches the S 11 response, since little attention has been paid to alternative activators.

Although data relevant to the photoelectric spectral response of $SbCs_3$ have been published by several observers, there are various discrepancies and errors among these data, and the factors which determine the spectral response do not appear to have been fully appreciated. In the present study measurements have been made of the relative photoelectric quantum efficiency $\eta(\lambda)$, over a range of wavelength λ from 200 mμ to 650 mμ, of the $SbCs_3$ cathodes of three commercial photomultipliers with Pyrex glass and fused quartz windows. In conjunction with other data, these have enabled the factors influencing the spectral response to be determined.

2. Experimental

The exit beam from a Unicam SP 500 spectrophotometer provided a monochromatic (bandwidth $<1\ m\mu$) light flux, a hydrogen discharge lamp being used as a source for $\lambda = 200$–370 mμ, and a tungsten filament lamp for $\lambda = 350$–600 mμ. In order to measure the output from the monochromator a 2·5 cm cubical cell of Spectrosil (Thermal Syndicate Ltd.) containing an integrating fluorescent solution was placed adjacent to the exit slit, and observed directly by the photomultiplier, which was mounted with its axis along the beam direction, and operated from a stabilized power supply. The cathode aperture was limited to 1 cm diameter to exclude any scattered light. The photomultiplier output current $I_0(\lambda)$, which did not normally exceed 3 μA and was stable to within 3 nA, was measured, the appropriate small correction being made for the dark current. $I_0(\lambda)$, which is proportional to the relative quantum intensity of the exit beam, was observed as a function of λ, using appropriate integrating solutions. The direct response current $I(\lambda)$ of the photomultiplier was similarly measured as a function of λ, with the Spectrosil cell, empty and clean, mounted in its previous position. The relative photoelectric quantum efficiency $\eta(\lambda)$ of the system was evaluated from the ratio

$$(\lambda) = I(\lambda)/I_0(\lambda). \qquad (1)$$

To function satisfactorily as an integrating solution, a solution must have a constant fluorescence quantum efficiency over the range of λ for which it is used, its absorption coefficient must be sufficiently high that the incident light is absorbed in a thin surface layer so that the optical geometry is substantially independent of λ, and it must be resistant to chemical or photo-induced changes. For $\lambda = 200$–370 mμ a 10^{-2} M solution of 1-dimethyl-aminonaphthalene 7 sodium sulphonate in distilled water was used, as proposed by Teale and Weber (1957) and further tested by Birks and Kuchela (1961), and this gave reliable and consistent results. For $\lambda = 350$–650 mμ a 5×10^{-3} M solution of Rhodamine B in ethylene glycol was used. The observed variation of $I_0(\lambda)$ with λ using this solution agreed closely with that

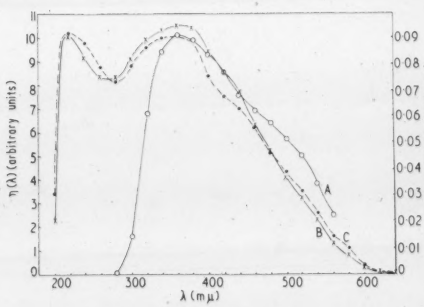


Fig. 1. Relative ($\eta(\lambda)$) and technical ($\eta_t(\lambda)$) photoelectric quantum efficiences plotted against wavelength λ. Curve A, 9514B (Pyrex, 44 mm cathode); curve B, 6255S (quartz, 44 mm cathode); curve C, 6256B (quartz, 10 mm cathode).

derived from the colour temperature of the tungsten-filament lamp.

Three integrating solutions proposed by Weber and Teale (1958), namely 10^{-2} M eosin in 0·1 N NaOH, 10^{-2} M fluorescein in 0·1 N NaOH, and 10^{-2} M fluorescein in N HCl, proved unsuitable since their quantum yield remained constant only over a limited range of λ, probably due to dimerization effects.

The measurements of $\eta(\lambda)$ against λ for three typical E.M.I. photomultipliers are plotted in Fig. 1. Each of the tubes has a thin $SbCs_3$ cathode, with a nominal luminous sensitivity of 50 μA/lm. In type 9514 B the cathode is of 44 mm diameter and is deposited on a 1·5 mm thick Pyrex glass window. In types 6255 S and 6256 B the cathodes are of 44 mm and 10 mm diameter respectively and are deposited on 1·5 mm thick fused quartz windows.

Within the limits of experimental accuracy, $\eta(\lambda)$ was observed to be independent of the potential V applied to the photomultiplier, for $V = 900$–1800 v, which corresponds to the normal range of operating potentials.

3. Analysis of results

The technical photoelectric quantum efficiency $\eta_t(\lambda)$ of a window-cathode system is defined as the fraction of photons of wavelength λ incident on the window which lead to the emission of an electron from the cathode into the multiplier dynode system. $\eta_t(\lambda)$ is determined by three factors:

(i) the transmittance of the window $W(\lambda, t)$, which depends on the transmission spectrum of the window material and on its thickness t;

(ii) the absorbance of the cathode $A(\lambda, d)$, which depends on the absorption spectrum of the cathode material and on its thickness d; and

(iii) the cathode photoelectric quantum efficiency $Q(\lambda, d, E)$, defined as the fraction of photons of wavelength λ absorbed by the cathode which lead to the emission of an electron.

Since Q depends on the ability of a photoelectron to escape from the cathode, it will be a function of d and of the electric field E at the cathode surface. In discussing the present results, E will not be considered, since the observed independence of $\eta(\lambda)$ and V indicates that Q is insensitive to E at the low fields (~50 v/cm) employed.

W determines the fraction of incident photons that penetrate to the cathode, A the fraction of these which are absorbed by the cathode, and Q the fraction of the latter which lead to electron emission, so that

$$\eta_t(\lambda) = W(\lambda, t)A(\lambda, d)Q(\lambda, d, E). \qquad (2)$$

$\eta(\lambda)$, the quantity observed, is proportional to $\eta_t(\lambda)$.

The transmittance spectra, $W(\lambda)$ against λ, of Pyrex glass and fused quartz for $t = 1{\cdot}5$ mm are plotted in Fig. 2. Comparison with the observations of $\eta(\lambda)$ in the short-wave cut-off regions (Fig. 1) shows that in these regions $\eta(\lambda)$ is simply determined by $W(\lambda)$, while AQ remains substantially constant. At longer wavelengths $W(\lambda)$ is independent of λ and does not therefore influence $\eta(\lambda)$.

Spicer (1958) has presented mean data, from the work of previous observers, on the optical absorption coefficient of $SbCs_3$ at $\lambda = 350$–700 mμ, and these are plotted in Fig. 3, curve A. Spicer has also observed the absolute photoemissive yield AQ of $SbCs_3$ cathodes of different thickness d at $\lambda = 320$–650 mμ. Our results agree closely with his for $d = 200$ Å, and this value has therefore been taken as the thickness of the cathodes studied. Comparison with Spicer's data also enables our values of $\eta(\lambda)$ to be normalized to absolute values (electrons/photon) of $\eta_t(\lambda)$. These values are shown on the right-hand ordinate scale of Fig. 1.

For $d = 200$ Å, $A(\lambda)$ has been evaluated from the absorption spectrum (Fig. 3) for $\lambda = 350$–650 mμ, and the values of $Q(\lambda)$ have been obtained from Fig. 1 and Eqn (2). Values

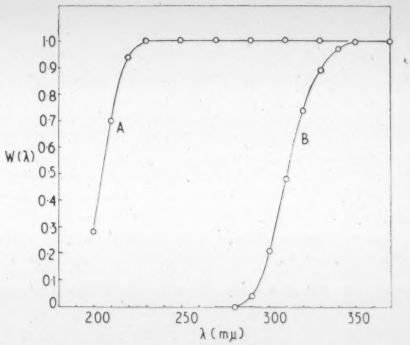


Fig. 2. Transmittance $W(\lambda)$ plotted against wavelength λ for $t = 1{\cdot}5$ mm of A, fused quartz window, B, Pyrex glass window.

of $Q(\lambda)$ for $d = 400$ Å have been similarly evaluated from Spicer's results, and $Q(\lambda)$ is plotted against λ in Fig. 4, curves B and C. It is found that Q is practically independent of λ for $\lambda = 390$–320 mμ.

On the assumption that Q remains substantially independent of λ for $\lambda = 350$–220 mμ, the values of $A(\lambda)$ for

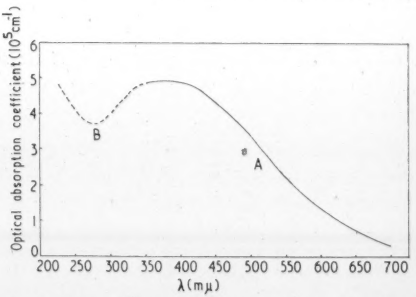


Fig. 3. Absorption spectra of $SbCs_3$. Curve A, $\lambda = 350$–700 mμ (Spicer 1958); curve B, $\lambda = 220$–350 mμ (from present results).

$d = 200$ Å have been evaluated from Fig. 1 and Eqn (2). The absorption spectrum of $SbCs_3$ for $\lambda = 350$–220 mμ thus obtained is shown in Fig. 3, curve B.

Spicer has measured AQ against λ for $d = 200$, 400, 800 and 1100 Å. His values of AQ against d at $\lambda = 350$ mμ, the corresponding values of A obtained from the absorption spectrum (Fig. 3), and the derived values of Q are plotted in Fig. 5.

The curve of Q against d has been extrapolated to Q_0 at $d = 0$, where

$$Q(\lambda, d, E) = Q_0(\lambda)f(\lambda, d, E). \qquad (3)$$

$f(\lambda, d, E)$ represents the electron 'escape probability' which decreases rapidly, from unity at $d = 0$, with increasing d,

and $Q_0(\lambda)$ is the absolute photoelectric efficiency of the cathode material. Spicer's data have been analysed and extrapolated to $d = 0$ in a similar manner for $\lambda = 320$–$650\,m\mu$, and the values of Q_0 thus obtained are plotted against λ in Fig. 4, curve A.

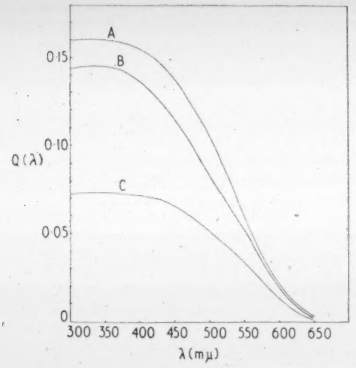


Fig. 4. Cathode photoelectric quantum efficiency $Q(\lambda)$ plotted against wavelength. Curve A, $Q_0(\lambda)$, $d = 0$ (extrapolated); curve B, $Q(\lambda)$, $d = 200$ Å; curve C, $Q(\lambda)$, $d = 400$ Å.

4. Discussion

4.1. *Comparison with other data*

The $SbCs_3$ cathode deposited on Pyrex glass has the response designated S 11. The spectral response data for this type of cathode published by the manufacturers differ markedly from the present results. In the original E.M.I. data sheets (up to 1958) the response was plotted in terms of relative sensitivity $S(\lambda)$, for equal energies of incident light at each wavelength, so that $S(\lambda) \propto \lambda\eta(\lambda)$. $\eta(\lambda)$ from the E.M.I. data is compared with the present results in the Table.

S 11 *spectral response data*

λ ($m\mu$)	$\tau(\lambda)$ (*present results*)	$\tau(\lambda)$ (*E.M.I. data*)
640	0	0·05
600	0·07	0·14
560	0·30	0·41
520	0·50	0·70
450	0·65	0·85
440	0·76	0·96
400	0·96	0·93
350	0·98	0·69
360	1·00	0·50
340	0·94	0·33
320	0·69	0·13
300	0·16	0

While the difference from $\lambda = 640$–$400\,m\mu$ may be partly due to differences between the photocathode thicknesses considered, the E.M.I. data at $\lambda < 400\,m\mu$ are almost certainly incorrect, since $\eta(\lambda)$ in this region has been shown to be determined by $W(\lambda)$ (Fig. 2). This spectral region is of particular importance in organic scintillator design, since many organic fluors emit in this region.

Sharpe and Thomson (1958) have published similar response data to the manufacturers, except that the ordinate of their graphs is marked 'quantum efficiency' instead of $S(\lambda)$, which is the quantity plotted. This error is introduced into more recent issues of the manufacturers' data sheets (1959–60) in which curves similar to their original ones are given, but with the ordinate marked 'quantum efficiency'. The error is also reproduced by Braddick (1960), quoting data from the same source.

The $SbCs_3$ cathode deposited on fused quartz has the response designated S 13. The measurements on the S 11 and S 13 responses (Fig. 1) are in excellent agreement with those on R.C.A. photomultipliers with similar cathodes (Dunkelman and Lock 1951, Engstrom, Stoudenheimer and Glover 1952, Maron 1955, Swank, Buck, Hayes and Ott 1958, Engstrom and Matheson 1960), when allowance is made for slight differences in cathode thickness and uniformity between individual tubes, which modify the response curves mainly in the region from $\lambda = 400$–$650\,m\mu$. Slight differences of this nature are seen in Fig. 1 for the three tubes studied, and Engstrom, Stoudenheimer and Glover (1952) have reported on the extreme variations observed in the production testing of R.C.A. photomultipliers.

The cut-off of the S 13 response at short wavelengths is determined by the transmittance spectrum of the window. Dunkelman and Lock (1951) have observed that $\eta(\lambda)$ is reduced to 10% of its 220 $m\mu$ value at $\lambda = 185\,m\mu$ for an R.C.A. 1 P28 photomultiplier with a thin Corning 9741 high silica glass envelope, and at $\lambda = 170\,m\mu$ for a similar tube with a quartz envelope. The present results indicate that the corresponding cut-off for E.M.I. quartz-windowed tubes is at $\lambda \simeq 195\,m\mu$, the higher value being due to the increased window thickness and the inferior transparency of the quartz used. The statement by Sharpe and Thomson (1958) that these tubes cut off at $\lambda = 160\,m\mu$ is incorrect.

Most observers have presented their data as $S(\lambda)$ against λ, rather than as $\eta(\lambda)$ against λ. The latter is to be preferred since it is directly related to the performance of the tube as a scintillation detector. In particular the latter mode of presentation does not obscure the simple but important feature, apparently not noticed hitherto, that the photoelectric quantum efficiency of a quartz-windowed $SbCs_3$ cathode is nearly independent of λ for $\lambda = 220$–$400\,m\mu$.

The luminous sensitivity, expressed in μA/lumen, which is commonly used by manufacturers in rating tubes, bears no relation to the use of a photomultiplier for scintillation counting, since it ignores the ultra-violet sensitivity and includes the red sensitivity.

4.2. *The spectral response*

In the short-wave cut-off region the spectral response of the cathode-window system is determined by the transmittance of the window. The ultra-violet cut-off could be extended to shorter wavelengths by the use of thinner windows or more transparent materials, e.g. high purity fused silica, lithium fluoride, etc. The response of glass-windowed tubes at $\lambda \leqslant 350\,m\mu$ will be affected by the transparency of the glass used.

The long-wave response is determined fundamentally by $Q_0(\lambda)$ (Fig. 4) and is influenced by $A(\lambda, d)$ and $f(\lambda, d, E)$. A and f increase and decrease respectively with d, so that the cathode peak efficiency AQ at $\lambda = 350\,m\mu$ is a maximum at $d = 150$–250 Å (Fig. 5). An increase in d will tend to shift the response curve towards longer λ, due to an increase in A, while it will reduce the technical efficiency. Hence cathodes with a peak response at $\lambda > 400\,m\mu$, due to an increase in d, will have a reduced efficiency and probably a higher thermionic emission than those with $d = 200$ Å. It appears from the spectral response data given by Engstrom, Stoudenheimer

and Glover (1952) on R.C.A. 5819 tubes, that the average tube of this type, in 1952, with a peak efficiency at $\lambda = 460$ mμ departed considerably from optimum performance. On the other hand the responses of the three E.M.I. tubes tested appear to correspond to the optimum value of $d = 200$ Å, evaluated from the R.C.A. Laboratories data of Spicer (1958).

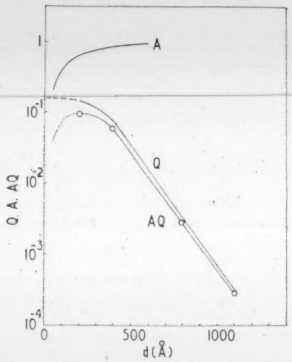


Fig. 5. Technical efficiency AQ, absorbance A, and cathode efficiency Q plotted against cathode thickness d. ($\lambda = 350$ mμ.)

One method of increasing AQ is to deposit the $SbCs_3$ cathode on a metal reflector, thus doubling the effective 'optical' value of d and increasing A, without much reducing the electron escape probability f. From Fig. 5 it is estimated that for $d = 150$ Å, $\lambda = 350$ mμ, A and AQ for such a solid cathode will be about 50% greater than for a similar semi-transparent cathode. The effect of the increased A will be to increase the spectral response at longer wavelengths, apart from increasing the technical efficiency. Solid cathodes are used in the R.C.A. 1P21 and 1P28 tubes and similar types. These have the S4 and S5 responses, which are of the form predicted. Although the solid-cathode tube is not to be recommended for scintillation spectrometry where uniform light collection is required, it is to be preferred to the end-window tube for applications concerned with the detection of weak collimated light fluxes, because of its higher sensitivity.

An alternative possible method of increasing Q is to increase the electric field E applied at the photocathode surface, thereby increasing the electron escape probability $f(\lambda, d, E)$ towards unity. An increase in E will probably also increase the dark noise. The limit to the value of Q obtainable will normally be Q_0, unless internal secondary emission occurs leading to electron multiplication within the cathode. The elementary condition for the latter effect is that the mean free path of the incident photon should exceed the mean free path of the primary electron for secondary emission. Spicer has observed for $d = 800$ Å and 1100 Å that the curves of AQ against λ have unusual maxima at $\lambda \simeq 560$ mμ, and this phenomenon, not previously explained, may be due to internal secondary emission which may occur due to the high d and relatively low optical absorption. The effect merits further study and may be of some practical interest in photoelectric devices, provided dark noise does not prove excessive.

$Q_0(\lambda)$ is the absolute photoelectric efficiency which is to be compared with theoretical analyses of the photoelectric effect, the form of the threshold curve being influenced by the Fermi–Dirac distribution. Previous observers have tended to consider η_t or AQ as equivalent to Q_0 in such analyses. It is of interest to note the parallelism between the curves of optical absorption coefficient against λ (Fig. 3) and of Q_0 against λ (Fig. 4), the latter being displaced towards the ultra-violet relative to the former by about 0·4 ev. A close association between the electronic transitions responsible for optical absorption and photoelectric emission is thus indicated.

4.3. *Scintillation counting*

As pointed out in the introduction, the S 11 response has strongly influenced the design and development of scintillator systems. The virtues of the S 13 response, i.e. $\eta_t > 0{\cdot}8(\eta_t)_{max}$ for $\lambda = 210$–430 mμ, compared with $\lambda = 330$–430 mμ, for the S 11 response have been largely ignored.

The effect of the use of quartz-windowed photomultipliers on the design of organic scintillator systems has been discussed elsewhere (Birks 1961, Paper NE/74, I.A.E.A. Conference on Nuclear Electronics, Belgrade (May), to be published). Inorganic scintillators similarly merit reconsideration, with freedom from the restriction of an emission spectrum which matches the S 11 response. Although thallium is an efficient activator of alkali halide crystals like NaI, CsI and KI, it is possible that other activators may be found which are equally or even more efficient, and may also yield systems with reduced scintillation decay times. A programme of work is being planned to explore these possibilities. Similarly in studies currently in progress here on gas scintillators it is hoped to eliminate the need for external wavelength-shifters by the use of photomultiplier windows transparent to the ultra-violet emission of the gas.

4.4. *Tri-alkali cathodes*

The photoelectric efficiency $\eta_t(\lambda)$ of the tri-alkali (SbM_3, where M is a mixture of two or more alkali metals, such as Na and K) cathode is higher than that of $SbCs_3$ (Sommer 1955). Similar studies are therefore being made of quartz-windowed tri-alkali cathode photomultipliers.

Acknowledgments

One of us (I. H. M.) wishes to acknowledge the tenure of a D.S.I.R. Research Studentship.

References

Birks, J. B., and Kuchela, K. N., 1961, *Proc. Phys. Soc.*, **77**, 1083.

Braddick, H. J. J., 1960, *Rep. Prog. Phys.*, **23**, 154.

Dunkelman, L., and Lock, C., 1951, *J. Opt. Soc. Amer.*, **41**, 802.

Engstrom, R. W., and Matheson, R. M., 1960, *Instn Radio Engrs Trans Nucl. Sci.*, **NS–7**, 52.

Engstrom, R. W., Stoudenheimer, R. G., and Glover, A. M., 1952, *Nucleonics*, **10** (4), 58.

Maron, H. S., 1955, *J. Opt. Soc. Amer.*, **45**, 12.

Sharpe, J., and Thomson, E. E., 1958, *Atoms for Peace Conference, Geneva.* (CP 154, E.M.I. Electronics Ltd.).

Sommer, W., 1955, *Rev. Sci. Instrum.*, **26**, 725.

Spicer, W. E., 1958, *Phys. Rev.*, **112**, 114; *R.C.A. Rev.*, **19**, 555.

Swank, R. K., Buck, W. L., Hayes, F. N., and Ott, D. C., 1958, *Rev. Sci. Instrum.*, **29**, 279.

Teale, F. W. J., and Weber, G., 1957, *Biochem. J.*, **65**, 476.

Weber, G., and Teale, F. W. J., 1958, *Trans Faraday Soc.*, **54**, 640.

Average kinetic energy of diffusing particles in discharge and electron tubes

There is some confusion in the literature of gas discharge tubes and electron tubes regarding the mean kinetic energy ϵ carried by particles with a Maxwellian velocity distribution flowing under a concentration gradient or small field. Most authors take this to be $2kT$, whereas the detailed kinetic theory of transport processes, as expounded by Chapman and Cowling (1939) and Hirschfelder, Curtiss and Bird (1954), shows it to be $\frac{5}{2}kT$. The author has traced this custom back to the classical review of Compton and Langmuir (1930)—it may well be of earlier origin. For many kinds of gas discharge the difference is of little importance since the energy involved is relatively small, but it becomes appreciable in certain cases.

One case of particular interest recently is that of 'thermionic generators', in which electrons flow from a hot cathode to a cooler anode across a region of small electrical field. Treatments by Hernqvist, Kanefsky and Norman (1958), Houston (1959), Nottingham (1959) and others, all take this energy to be $2\,kT$, whilst Lewis and Reitz (1960) use $2\,kT$ and $\frac{5}{2}\,kT$ in different places. It is therefore of interest to investigate this discrepancy.

The mean kinetic energy per particle ϵ carried across a surface normal to the z direction by a group of particles in a gas with a velocity distribution function $f(V)$ is

$$\epsilon = \frac{m}{2} \cdot \frac{\int V^2 V_z f(V) d\mathbf{V}}{\int V_z f(V) d\mathbf{V}}, \tag{1}$$

where the integrals are to be carried out over all directions in velocity space, including both hemispheres about the z direction. In the case of an equilibrium Maxwellian velocity distribution

$$f(V) = f^0(V) = A \exp(-\beta V^2), \tag{2}$$

where $\beta = m/2kT$ and $A = (\beta/\pi)^{3/2}$, and evaluation of (1) gives $\epsilon = 0$, by symmetry. The result $\epsilon = 2\,kT$ is obtained by considering only those particles with a positive velocity component V_z, i.e. by evaluating the integrals in (1) over one hemisphere only. This gives, on using the symmetry about the axis of flow,

$$\epsilon = \frac{m}{2} \cdot \frac{\int_0^\infty V^5 \exp(-\beta V^2) dV}{\int_0^\infty V^3 \exp(-\beta V^2) dV} = \frac{m}{2} \cdot \frac{1/\beta^3}{1/2\beta^2} = 2kT. \tag{3}$$

Although this result may be applicable to the effusion of particles from a large body of gas through a small hole, it is not valid for particles flowing across a surface within the body of the gas. In this case the integrals in (1) must be evaluated over all directions in both hemispheres. Since the equilibrium Maxwellian distribution gives, as it should, a zero result, it is necessary to consider the perturbation in the distribution corresponding to the flow. This is done in the general theory given by Chapman and Cowling (1939) for a single component gas and binary mixtures, by Hirschfelder *et al.* (1954) for multicomponent mixtures, and by Marshall (1960) for fully ionized gases, yielding in each case $\epsilon = \frac{5}{2}kT$. These all evaluate the theory in a very general form, however, using the elegant Enskog perturbation procedure, and it is difficult to trace the origin of the factor $\frac{5}{2}$. The following considerations illustrate how it arises.

When there is a flow of particles the Maxwellian distribution (2) is modified thus

$$f(V) = f^1(V) = f^0(V)\{1 + \phi(V)\}, \tag{4}$$

where the perturbation $\phi(V)$ is not symmetrical in all directions as $f^0(V)$, but only about the axis of flow. A very satisfactory approximation is

$$f^1(V) = f^0(V)[1 + \phi V_z], \tag{5}$$

where ϕ is a constant; terms involving higher powers of V_z, or the other components of V, have only a very small effect (Present 1958). Using (5) in (1), the first term vanishes by symmetry and the remainder reduces to

$$\epsilon = \frac{m}{2} \cdot \frac{\int_0^\infty V^6 \exp(-\beta V^2) dV}{\int_0^\infty V^4 \exp(-\beta V^2) dV} = \frac{m}{2} \cdot \frac{15}{16\beta^3}\left(\frac{\pi}{\beta}\right)^{1/2} \Big/ \frac{3}{8\beta^2}\left(\frac{\pi}{\beta}\right)^{1/2} = \frac{5}{2}kT, \tag{6}$$

which may be compared with (3). Thus the factor $\frac{5}{2}$ is seen to arise as the direct result of the perturbation in the velocity distribution.

The general theories cited show that this is valid under all ordinary conditions of flow. Although they only consider ordinary gas mixtures, and fully ionized gases in which both ions and electrons are at the same temperature, it is evident from the above demonstration that the result is not dependent on the detailed mechanism of each individual case and it is almost certainly applicable also to partially ionized gases and to plasmas in which the electron and gas temperatures are different.

Now consider a flow of electrons between a uniform plasma region and an electrode. There is normally a space-charge sheath immediately in front of the electrode, either accelerating or retarding to the electron flow, and next to this a region of disturbed plasma, in which the electron velocity distribution differs from (5) by a greater or lesser degree. (For example, if there is a net flow of electrons from the plasma into a sheath which accelerates them to a positive electrode, there will be no electrons at all moving away from the electrode at the very edge of the sheath; the distribution (5) will only be established gradually, across the disturbed region.) In considering the energy balance of the plasma, the flow of energy is required across a surface separating the bulk plasma, in which the distribution has the form (5), and this disturbed region. The average electron kinetic energy contribution to this is, of course, $\frac{5}{2}kT_e$, where T_e is the electron temperature, as has been shown; whether the net electron flow is to or from the electrode, and whether the sheath is accelerating or retarding to this flow, is immaterial.

In order to determine the energy actually reaching an electrode to which, for example, electrons are flowing, if this

should be required, a further energy balance across the sheath plus disturbed plasma region may be needed. If there is no means of energy exchange in this region between the electrons and other particles (ions or atoms), the electron contribution is simply the algebraic sum of the total potential drop across this region and $\frac{5}{2}kT_e$. In cases where appreciable energy is exchanged in collisions or electron-ion interaction processes, however, an overall energy balance must be used, including the ion energy flux and thermal conduction fluxes across the dividing surface.

It is essential to remember that T_e, in these formulae, is the electron temperature in the bulk plasma. If the disturbed distribution at the very edge of the sheath can still be characterized by an electron temperature, T_s say, this may differ from T_e; for example, if it is assumed half-Maxwellian, there being no electrons returning from the electrode with speeds corresponding to those passing to it, then the mean energy of the latter will be $2kT_s$, in accordance with (3). Thus, again provided there is no means of energy exchange with other particles, $2kT_s - eU_s = \frac{5}{2}kT_e - eU_p$ in this case, where U_s is the potential at the sheath edge and U_p that in the bulk plasma, a relation relevant to the determination of electron temperatures by Langmuir probes.

The author thanks Dr. C. A. G. Enstone, with whom he has discussed this subject.

Research Laboratory,
Associated Electrical Industries (Rugby) Ltd.,
Rugby, Warwickshire.

M. A. CAYLESS

14*th June* 1961, *in revised form* 12*th July* 1961

CHAPMAN, S., and COWLING, T. G., 1939, *The Mathematical Theory of Non-Uniform Gases* (Cambridge: University Press).
COMPTON, K. T., and LANGMUIR, I., 1930, *Rev. Mod. Phys.*, **2**, 123.
HERNQVIST, K. G., KANEFSKY, M., and NORMAN, F. H., 1958, *R.C.A. Review*, **19**, 244.
HIRSCHFELDER, J. O., CURTISS, C. F., and BIRD, R. B., 1954, *Molecular Theory of Gases and Liquids* (New York: John Wiley).
HOUSTON, J. M., 1959, *J. Appl. Phys.*, **30**, 481.
LEWIS, H. W., and REITZ, J. R., 1960, *J. Appl. Phys.*, **31**, 723.
MARSHALL, W., 1960, *A.E.R.E. Reports* T/R 2247, 2352 and 2419 (London: H.M.S.O.).
NOTTINGHAM, W. B., 1959, *J. Appl. Phys.*, **30**, 413.
PRESENT, R. D., 1958, *Kinetic Theory of Gases* (New York: McGraw-Hill).

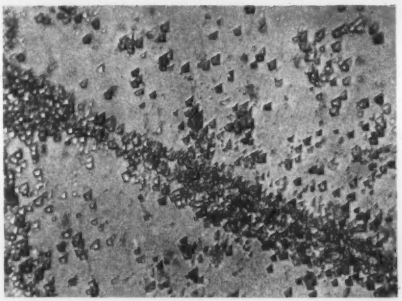
Dislocation etch pits in polished lead telluride

Although several methods have been described in the literature for etching lead telluride it has not previously been possible to polish and etch random sections of single crystal material satisfactorily. It is comparatively straightforward to etch freshly cleaved surfaces and there are two etchants available which produce pyramidal etch pits (Coates 1957†, Houston and Norr 1960). Unfortunately this technique suffers from the drawbacks that only (100) surfaces can be examined since at ambient temperatures lead telluride cleaves along these planes and in cleaving the crystals it is difficult to produce a surface which is not interrupted by cleavage steps and river patterns.

Attempts to etch mechanically polished surfaces were only partially successful. Lead telluride is so soft (VPN 30) that extreme care must be taken to produce a clean surface and moreover, there is the risk that the surface has been plastically deformed by the polishing action.

An electrolytic polishing technique (Schmidt 1961) has been slightly modified to suit the equipment available in the author's laboratory. The specimen is prepared by polishing on a 700 mesh carborundum powder and cleaning with an ultrasonic wash in carbon tetrachloride. It is preheated in a water bath or oven to about 70° C to reduce thermal shock during electro-polishing. The electrolyte consists of 100 ml. reagent grade 88% orthophosphoric acid and 10 g reagent chromic acid. The temperature must be held at 90–105° C. The specimen is made anodic with a tantalum strip cathode. A 6–20 v supply is required to produce a current density of about 3 A/cm² (a 12 v accumulator can be used). During polishing the specimen must be continuously agitated to prevent the deleterious effects of gassing. A time of 4–6 min is sufficient to produce a mirror-like finish though macro cracks often become evident. These may be caused through thermal shock during the electrolytic polishing or through growth fissures which are often present.

Upon removal from the electrolyte the specimen must be dipped immediately into alcohol at 70° C to suppress the formation of tenacious oxide layers. A white scum which forms can be wiped off with cotton wool and the specimen washed in a jet of cold water.



Electrolytically polished lead telluride etched with KOH to reveal dislocation etchpits on the (115) face (× 560)

Chemical etching must be performed whilst the specimen is still fresh. The etchant (Coates 1957†) consists of 10 vol KOH (1 g/cm²), 10 vol glycol and 1 vol H_2O_2. Oxidation is avoided and the etching is considerably improved if the whole apparatus is put into an inert atmosphere during polishing and etching. Well defined dislocation etch pits are readily revealed and dislocation densities, sub grain structure and deformation patterns may be examined.

Associated Electrical Industries Ltd.,
Semiconductor Materials Research Group,
Rugby, Warwickshire.

G. P. TILLY

28*th June* 1961

HOUSTON, B. B., and NORR, M. K., 1960, *J. Appl. Phys.*, **31**, 615.
SCHMIDT, P. H., 1961, *J. Electrochem. Soc.*, **108**, 104.

† D. G. Coates, R.R.E. Memorandum No. 1374 (1957).

Electron-optical conditions at Venetian blind type dynodes and their effects in photomultipliers

In the last few years we published a number of papers (Görlich, Krohs, Pohl and Schmidt 1957, Görlich, Krohs, Pohl and Wolf 1958, Görlich, Krohs, Pohl, Reichel and Schmidt 1958) which were concerned with the problem of designing electron-optically efficient photomultipliers for photometry as well as scintillation spectroscopy. However, we frequently experienced difficulties in the choice of the most suitable electron optics possible for the dynode stages; in particular we were interested in the question as to whether the employment of Venetian blind type dynodes would be of special advantage. Contrary to several research workers (Sommer and Turk 1950, Eckart 1958) who took the view that Venetian blind type dynodes in photomultipliers belong to the so-called unfocused systems, Jennings and Misso (1956) reported that a structure could be used "by taking advantage of the focusing property of the slat structure" and this incited us to tackle this problem once more.

Investigations of the electron-optical conditions of Venetian blind type dynodes by means of the electrolytic trough proved that only those secondary electrons which are emitted from a certain part of the slats facing the following dynode get into the directly extracting field of the following stage (cf. Fig. 1), while the electrons originating from other parts

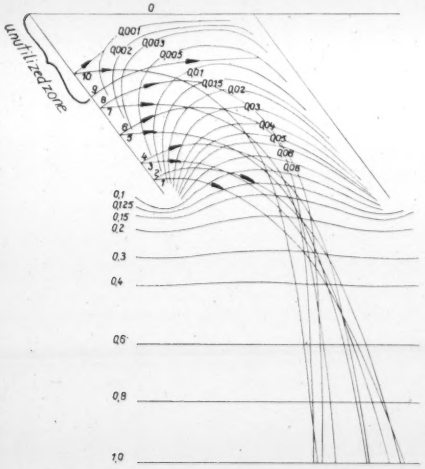


Fig. 1. Potential relief and electron paths at different angles of emergence in conventional Venetian blind type dynodes.

reach the following dynode only partly or indirectly (reflections from the slat or the potential relief above the dynode).

From a variety of all possible slat angles and proportions of covering meshes an optimal efficiency is obtained by a dynode structure as shown in Fig. 2, where only a wire is stretched parallel to and between the slats in the plane of the dynode, the wire being at the same potential as the slats.

As a result the field of the preceding dynode penetrates this 'grid' and makes the secondary electrons effective in that it constrains the electrons in the upper region of the slats to travel in the direction of the following dynode. In all the other cases the secondary electrons released in the upper region of the slats are directed on to the following dynode

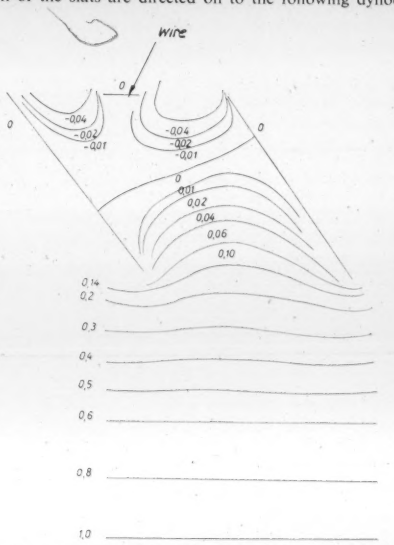


Fig. 2. Improved Venetian blind type dynode with enhanced electron-optical efficiency.

only with great delay, or only partly, or are possibly not attracted at all. This is demonstrated by the oscillogram shown in Fig. 3; here the photocathode of a Venetian blind type photomultiplier has been scanned by a light spot at right angles to the slat direction.

The slatted dynode was spaced only 21 mm from the plane

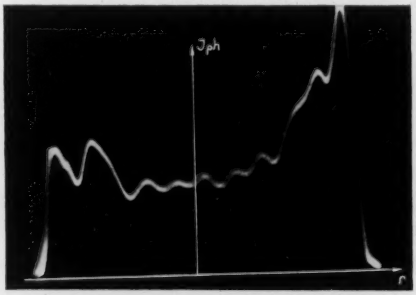


Fig. 3. Oscillogram of a Venetian blind type photomultiplier; the photocathode was scanned by a light spot at right angles to the dynode slats arranged in Venetian blind fashion. Photocurrent I_{ph} amplified by the first dynode as a function of releasing point r on the photocathode.

photocathode and the potential was chosen sufficiently high (450 v) for the electron beam released by the light spot to be about 2·5 mm in diameter on the dynode; thus the real structure of the dynodes which had a slat distance of approximately 3 mm was blurred by a factor of 3–5. The electron-optical efficiency of the dynodes may therefore at certain points drop to a few per cent.

The slats make an angle of about 30–40° with the plane of the dynode in the case of the multiplier investigated, Type V12J1 (Eckart 1958). The slat width is equal to the slat distance in the plane of the dynode. In the case of the Venetian blind type secondary electron multiplier, Type EMI (20th Century Electronics Ltd.) an angle of 35–40° is preferred, whereas in the Soviet secondary electron multipliers, Types FEU 11–16 preference is given to an angle of 59°.

From these experimental results two conclusions may be drawn for the development of photomultipliers:

(1) We are fairly confident that the nuclear-spectrometric resolution can be improved from a half-width of 7·5% to a half-width of 6% or less, if a secondary emission system were used that would be more ideal than the blind type dynode system.

(2) Venetian blind type multipliers are of no use when measurements are involved calling for a short rise time.

VEB Carl Zeiss Jena,
Jena,
Germany.

P. Görlich
A. Krohs
H.-J. Pohl

2nd May 1961, *in revised form* 12*th July* 1961

Eckart, F., 1958, *Exper. Tech. der Phys.*, **6**, 62.
Görlich, P., Krohs, A., Pohl, H.-J., Reichel, R., and Schmidt, L., 1958, *Z. Angew Phys.*, **10**, 303.
Görlich, P., Krohs, A., Pohl, H.-J., and Schmidt, L., 1957, *Exper. Tech. der Phys.*, **5**, 1.
Görlich, P., Krohs, A., Pohl, H.-J., and Wolf, G., 1958, *Acta Imeko*, **2**, 293.
Jennings, A. E., and Misso, C. E. F., 1956, *J. Sci. Instrum.*, **33**, 324.
Sommer, A., and Turk, W. E., 1950, *J. Sci. Instrum.*, **27**, 113.

Araldite used for model analysis of composite structures

Stress analysis in composite models consists of having contact joints between two materials of different elastic properties using a separate adhesive, and in certain cases this would not quite satisfy the idealized conditions expected in a joint.

Consequently a new technique is developed wherein Araldite casting resin D or B with and without filler is used to represent two materials of different elastic properties used in the model. The problem under investigation is stress determination in layered systems in dams and this has been undertaken in connection with one of the biggest hydro-electric projects in India.

The model representing the cross section of the dam is to consist of two materials of different elastic properties in layers one above the other, the ratio of their moduli being 1:1·5. In the case of a dam consisting of concrete and masonry this ratio is 1:2·5. It is proposed to determine the unknown self-equilibrating stresses that develop at and in the neighbourhood of the joint for hydrostatic and gravitational forces separately. This needs a joint between two materials of high tensile strength. In a homogeneous model, these correction or self-equilibrating stresses would be absent.

Araldite is used for the top portion of the dam model and for the lower portion Araldite with silica flour as filler which is of higher modulus of elasticity. The procedure is to cast the lower portion first and cast the upper one integral with the former.

Experiments were conducted in our laboratory to study the effect of two types of fillers on the Young's modulus of Araldite. The results are given in Tables 1 and 2. The maximum percentage of filler, used in each case in Tables 1 and 2, indicates the quantity of the filler, which when added to the mixture made it so viscous that it settled in the mould with considerable difficulty. Since a higher modulus of elasticity and a more homogeneous material could be achieved in the case of Araldite with silica flour as filler, it was used in our experimental models.

Table 1. *Araldite and silica flour*

% of filler	*E*
70%	$0{\cdot}41 \times 10^6$
100%	$0{\cdot}50 \times 10^6$
125%	$0{\cdot}65 \times 10^6$
150%	$0{\cdot}98 \times 10^6$

Table 2. *Araldite and aluminium powder*

% of filler	*E*
20%	$0{\cdot}41 \times 10^6$
30%	$0{\cdot}48 \times 10^6$
40%	$0{\cdot}62 \times 10^6$

Experiments were also conducted on the strength of the joint between the two materials and it was found to vary with

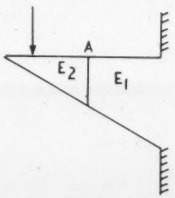


Fig. 1.

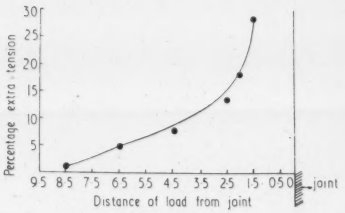


Fig. 2.

the percentage of filler added to the resin. The tensile strengths given in Table 3 is the maximum bending stress at which the joint failed.

Table 3. *Breaking strength of Araldite with silica flour*

% of filler	*Breaking strength*
150%	2220 lb/in^2
125%	2700 lb/in^2
100%	3200 lb/in^2

Also it was found that the initial stresses at the joint (which ideally should be absent for good results in photo-elasticity) were insignificant. In order to achieve high values of E_1/E_2, so that the results could be compared to the existing theoretical results (Silverman 1955), a plasticizer, viz., dibutyl phthalate was tried in different percentages, and it was found that the values of Young's moduli could be reduced to as low a value as about 900 lb/in^2. But it is not used at present due to the excessive creeping tendencies of the material.

Measurements were made with the aid of $\frac{1}{4}$ in. gauge length strain gauges (see Fig. 1), of the tensile stress at the joint, due to a concentrated force acting at a specific distance from the joint. The results were compared with that of a homogeneous model. Fig. 2 shows the relationship between the difference in the tensile stresses at A in composite and homogeneous models, with the distance of the load from the location of the joint. It is also interesting to note that the difference in tension between homogeneous and composite models decreases with the increase in distance between the concentrated load and the joints. The presence of excess tension in the composite model is comparable to what we can guess from the theoretical results of a wedge resting on a rigid foundation (Silverman 1955).

Detailed stress studies are being conducted with the aid of photo-elasticity and strain gauges for gravity and hydrostatic loading on both homogeneous and composite models.

I am deeply indebted to Professor N. S. Govinda Rao, for the keen interest and constant encouragement given by him during this work.

Civil Engineering Section,
Indian Institute of Science,
Bangalore,
India.

R. S. ALWAR
4*th July* 1961

I. K. SILVERMAN, 1955, *J. Appl. Mech.*, **22**, 123.

NOTES AND NEWS

Notes and comments

Annual General Meeting of The Institute of Physics and The Physical Society

At the first Annual General Meeting of the amalgamated Institute of Physics and The Physical Society held in London on 4th July 1961, the following were elected to office:

President: Sir John Cockcroft; *Vice-Presidents:* Dr. V. E. Cosslett, Professor R. W. Ditchburn, Dr. W. H. Taylor, and Dr. J. Topping; *Honorary Treasurer:* Dr. J. Taylor; *Honorary Secretary:* Dr. C. G. Wynne.

Eight ordinary members of Council were also elected. This, the first elected Council of the amalgamated body, is very nearly the same as the 'caretaker' Council which has been in office during the first few months of the amalgamated body.

The report for 1960 which was adopted at the meeting, records that The Physical Society and The Institute of Physics ceased to exist as separate corporate entities on 5th July 1960, but for convenience and completion of the records their activities before amalgamation are included in this report with those of the amalgamated body.

The report records the holding of many conferences, symposia and other meetings, many of which were arranged by the Branches and specialist subject Groups. Details about the four scientific periodicals and the other publications owned by the Institute and Society or for which it has been responsible are given in the report.

The report shows that at the end of the year the total membership in all grades was 8720. Of the 193 candidates who took papers in the Graduateship Examination, 63 were successful and of these 27 held university degrees and 35 the Higher National Certificate in Applied Physics. The number of candidates taking these examinations has continued to increase rapidly. In 1960, 46 technical colleges presented 194 candidates for the Ordinary National Certificate in Applied Physics and 29 colleges presented 355 candidates for the Higher National Certificate. During the year the Institute and Society issued a report entitled 'The Post-Graduate Training of Physicists in British Universities' which created a good deal of interest, being referred to with approval in the scientific, educational and lay press.

Among the professional matters referred to in the report is the publication of the fifth survey of salaries received by members of the professional grades.

In order to ensure that the work of the Benevolent Fund of The Institute of Physics should be continued uninterrupted by the amalgamation the Board of that Institute in its final meeting appointed the amalgamated body as a Trustee of the Fund. During the year, the report states, assistance was given in nine cases to members or their dependants who had met with misfortune.

Letters to the Editor

We should like to draw readers' attention to the publication of Letters to the Editor in this journal. We have arranged with our printers to publish letters received up to six weeks before the date of publication, provided they are reported upon favourably by our referees. For example, items for publication in the November issue of the journal should be received in the office not later than 20th September. Letters from the Continent will be published in French or German if necessary to avoid further delay. This also applies to the other two journals of The Institute of Physics and The Physical Society, i.e. the Proceedings of the Physical Society and the Journal of Scientific Instruments.

Notes for the Preparation of Papers for the *British Journal of Applied Physics*

The following instructions refer to the main points which have to be considered in the preparation of a paper for publication, and authors offering papers to the *British Journal of Applied Physics* for publication are asked to conform to these recommendations.

Manuscripts

Manuscripts should be typed in *double* spacing on paper not wider than 8 in. and not longer than 13 in. Only one side of the paper should be used, and a margin of 1–1½ in. should be left. As alterations in the text cannot be allowed once the paper is set up in type, authors should aim at absolute clarity of meaning and of typing, and should check the typescript carefully before submission. All manuscripts should be submitted in duplicate, and in addition to the fair copies, a set of small copies or prints of the diagrams (not larger than foolscap) must be attached to each MS. (This enables MSS. to be sent to two referees at the same time and so assists rapid publication.)

Abstract. An abstract is printed at the beginning of the paper immediately after the title, name(s) of author(s), and place of employment of author(s). Two extra copies of the 'title and abstract' page are required for the records.

Mathematics. It is not necessary to give detailed derivations of mathematical expressions and formulae in a published paper when the work is straightforward; it is quite sufficient to indicate the method of treatment and the final results.

References. In the text bibliographical references are made by giving the name of the author and the year of publication in brackets, e.g. (Jones 1942), and details are given in the last section, 'References', where the references are arranged in alphabetical order of authors' names and in date order for each author.

Drawings

Drawings should be in Indian ink on tracing cloth, tracing paper or white card, with lettering in soft or blue pencil; lettering of a size suitable for reduction will be inserted by our draughtsman. The drawings should in general be sufficiently large to allow of reduction in printing, and the lines should therefore be bold; the frame lines of graphs should be slightly finer than those of the plotted curves.

Full details are available in the Institute and Society's 'Notes for Authors', obtainable from the Editor and Deputy Secretary, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London S.W.7.

Papers for publication in the *British Journal of Applied Physics* should be sent to the Editor and Deputy Secretary at the same address.

British Journal of Applied Physics Editorial Board

Members of the Editorial Board for the British Journal of Applied Physics are as follows:

J. M. A. Lenihan, Ph.D., M.Sc., A.M.I.E.E., F.Inst.P. (*Chairman*).

Prof. R. W. Ditchburn, M.A., Ph.D., B.Sc., F.Inst.P. (*Ex-officio member*).

W. Hirst, Ph.D., B.Sc., F.Inst.P.

C. A. Hogarth, Ph.D., B.Sc., F.Inst.P.

J. W. Menter, M.A., Ph.D., F.Inst.P.

A. C. G. Menzies, M.A., D.Sc., F.Inst.P.

Prof. D. A. Wright, D.Sc., F.R.A.S., F.Inst.P.

Journal of Scientific Instruments

Contents of the September issue

PAPERS

Device designed to maintain a constant gas pressure in kinetic systems. By E. R. S. Winter.
Rotational vibroscope for the comparison of the torsional properties of thin fibres. By P. Nordon.
Production of light pulses of nanosecond rise time and duration by means of gas-discharge tubes. By Zs. Náray and P. Varga.
Ionization gauge power supply for use in a pulsed magnetic field. By G. A. Doran.
Smooth power control for high current furnaces. By B. Riley.
Response of an X-ray opposite-window monitor system to changes in tube conditions. By B. W. Delf.
Emission control for the omegatron-type mass spectrometer. By G. R. Giedd and G. C. Roberts.
Two simple methods for spot welding wires. By S. V. Radcliffe and J. S. White.
Magnetic crack detector. By P. H. Dawson.
Method for establishing a liquid column of graded density. By J. M. Jones.
Device for the accurate reproduction of liquid levels and some applications. By G. F. Hewitt and E. Lyall.

LABORATORY AND WORKSHOP NOTES

Ice point apparatus for hot climates. By T. D. Bansal.
Laboratory method for making small alumina crucibles. By I. Levanone.
Note on temperature measurement in high temperature X-ray powder cameras. By W. Johnson.
Device for stabilization of gas flow. By J. Novák and J. Janák.
The 'EEL' B.R.S. daylight photometer. By P. Petherbridge and W. M. Collins.
Seal between silica tubing and glass apparatus. By W. R. George and R. V. Winkle.
Simple electric pantograph. By G. J. Frisone.
Substitute for metal boats used in vacuum evaporation. By R. L. N. Sastry.
Objective astigmatism correction of the Elmiskop I electron microscope. By M. A. Silva.
Locking tweezers made from a hypodermic needle. By D. I. James and W. G. Newell.
Switch with floating preset differential. By W. Nisbet.
Cutting cam grooves by lathe. By A. L. Sims.
Switching by surface discharges. By J. S. T. Looms.
Construction of a cadmium-lined box for neutron experiments. By J. B. C. Brown and R. S. Hall.

NOTES AND NEWS

New instruments, materials and tools Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.
EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.
REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.
ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611-2.)
SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12*s.* 6*d.* each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.
CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

PRINTED IN GREAT BRITAIN BY UNWIN BROTHERS LIMITED, WOKING AND LONDON

INVITED ARTICLE

Lead zirconate–titanate piezoelectric ceramics

by A. E. CRAWFORD, M.Brit. I.R.E., Brush Crystal Co. Ltd., Hythe, Southampton

Paper presented at the Institute of Physics and the Physical Society Acoustics Group Meeting, 2nd March 1961

MS. received 2nd June 1961

Abstract

The disadvantages of ceramics based on barium titanate has limited many of the possible applications of this type of piezoelectric material. The Curie point is low and it is inefficient in electromechanical conversion. The discovery of the piezoelectric properties of the solid solution ceramics based on lead zirconate and lead titanate has stimulated increasing interest in transducer design. The paper describes the properties and characteristics of this new range of materials in terms of their ferroelectric and piezoelectric effects.

Introduction

The development and use of piezoelectric ceramic materials has been very rapid and covers a period of only about fifteen years. Before their discovery it was generally accepted that piezoelectricity was an attribute found only in single crystals having an intrinsic polar nature, such as quartz and rochelle salt. There had been some speculation about piezoelectric effects in oriented aggregates but the effects were small and uncertain.

In early 1940 the extraordinarily high dielectric constant in a barium titanate oxide ceramic was noted by the staff of an American industrial ceramics laboratory (Wainer 1946). Peculiar dielectric phenomena were found in this material during subsequent work, but it was left to R. B. Gray (1949) to discover that it could be made to exhibit substantial and practically useful piezoelectric effects.

The basic property of barium titanate and related compounds which enables them to behave as piezoelectrics is ferroelectricity. This is the ability of certain crystals of electrically polar structure to switch the direction of polarity under the influence of a strong electric field between several crystallographically equivalent directions, and to retain their new orientation after removal of the field.

In the crystal form of barium titanate the spontaneous electrical polarization may be directed positively or negatively along any one of the three cube axes. When barium titanate is prepared as a ceramic the individual crystallites have essentially random orientation and the polarity is distributed statistically among the several permissible directions and senses within each crystal. Regions of uniformly directed electric moment are known as domains. Most ferroelectrics lose their polarity and hence their domain structure at a characteristic temperature, the Curie point. A strong electric field such as 20 kv/cm applied to barium titanate ceramic, preferably at a temperature just below the Curie point of 120° c, changes the domain pattern and causes a preferential alignment of the polarity in directions including the smallest possible angle with the applied field. Before this polarizing process the ceramic as a whole is essentially isotropic and not piezoelectrically responsive. After polarizing there is an overall polarity and it is strongly piezoelectric.

The microphotograph in Fig. 1 clearly shows the domain patterns in barium titanate ceramic. The sample was deeply etched to bring out the boundaries, the average grain size being about 32 microns.

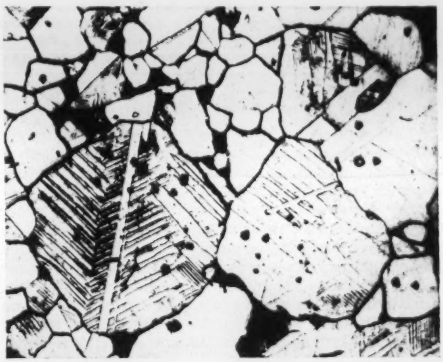


Fig. 1. Microphotograph of etched barium titanate ceramic.

Lead zirconate–titanate

While working at the U.S. National Bureau of Standards, B. Jaffe noted a Japanese report on the existence of two ferroelectric phases in the solid solution lead zirconate–titanate (Shirane and Suzuki 1952). This resulted in a study of the dielectric and piezoelectric properties of the system. Jaffe observed that substantial piezoelectric effects could be induced in compositions near the phase boundary separating regions of rhombohedral and tetragonal geometry (Jaffe, Roth and Marzullo 1954). It will be seen from Fig. 2 that piezoelectric activity is confined to a narrow region of solid solution ratios and there are pronounced maxima of piezoelectric and dielectric constants near the critical compositions. They are characterized by Curie points between 300 and 400° c and show larger piezoelectric effects than barium titanate. In analogy to the magnetostrictive case, these compositions may be stated to have low anisotropy energy and that is probably the reason why they can be polarized despite the higher level of ferroelectric crystal distortion.

The two constituents of the ceramic are worthy of close study as both possess unique electrical properties. Lead titanate is ferroelectric with a Curie temperature of about

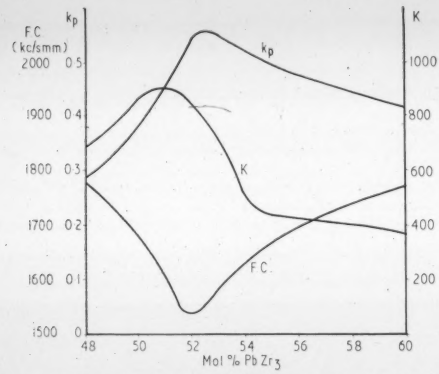


Fig. 2. Variations of parameters with varying ratios of lead zirconate and lead titanate.
F.C., frequency constant; k_p, planar coupling coefficient; K, relative dielectric constant.

490° C and there are no lower phase transformations such as are found in barium titanate. It possesses interesting dielectric properties and should show strong piezoelectric effects. However, it is very difficult to prepare as a ceramic and the required polarization fields at room temperature are very large. Increasing conductivity near the Curie point prohibits polarization at higher temperatures.

Lead zirconate also exhibits interesting dielectric properties characterized by a sharp rise to a relative dielectric constant of over 300 at the transition point of 200° C. The material can be considered as an antiferroelectric as below the transition temperature there is an absence of hysteresis and there is a linear relationship between field strength and polarization. However, at temperatures nearly approaching this value small hysteresis loops appear at each end of a linear trace. This demonstrates that large fields will cause a transition in the crystal from the antiferroelectric to the ferroelectric state.

The manufacture of lead zirconate–titanate presents certain difficulties owing to the low-temperature volatilization of the lead constituent. A number of methods of preparation have been described, a typical procedure being outlined by Kulcsar (1959a). PbO, ZrO_2 and TiO_2, in the correct ratios, are wet-mixed by ball milling, dried and then heated at 850° C for a few hours in a sealed crucible. The reacted material is then dry ground and a suitable binder added. The required shape is then pressed in a steel die at a pressure of approximately 10 000 lb/in^2. After being dried the pressings are fired in sealed alumina pots containing a $PbZrO_3$ pressed disk to provide a PbO atmosphere. A firing temperature of 1280 to 1290° C is used. The fired shapes are ground or lapped to dimensions and the requisite surfaces provided with fired-on silver electrodes.

Modified ceramics

The basic solid solution lead zirconate–titanate possesses a number of disadvantages for many piezoelectric applications, particularly a low dielectric constant and a marked temperature dependence. As with barium titanate, it is possible to modify many of the piezoelectric and electrical characteristics by certain additives, and this has been the subject of extensive studies in the United States of America, and this country (Kulcsar 1959 a,b, Crawford 1959, Gerson 1960). These additives need not enter into the solid solution and generally constitute only minor chemical variations. They result in considerable changes in properties such as dielectric constant, mechanical Q and ageing characteristics.

The most marked changes in performance are produced by the addition of small quantities of the oxides of tri or penta-valent elements such as lanthanum, neodynium or niobium and the practical exploitation of these additives is found in the commercially available material known as LZ-5 or PZT-5.

The major effects of these elements can be summarized as follows:

Ageing. All ferroelectric ceramics exhibit some ageing of their dielectric and mechanical properties during shelf life and also after drastic thermal or electrical excursions. The ageing of the modified ceramic is however almost an order of magnitude smaller than that of the plain ceramic. Figs 3 and 4 show the frequency constant, dielectric constant and

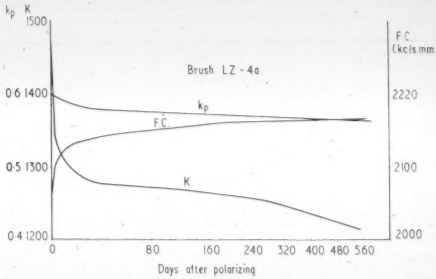


Fig. 3. Ageing effects in unmodified ceramic type LZ-4a.

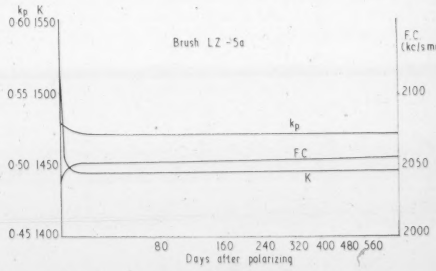


Fig. 4. Ageing effects in modified ceramic type LZ-5a.

planar coupling factor in terms of time after polarizing for unmodified and modified ceramic. It will be noted that the major changes in characteristics occur during the first week after polarizing.

Internal losses. The mechanical and electrical losses under low stress or field encountered in the modified material are considerably higher than in the unmodified ceramic. For

example, the dielectric loss angle tan δ is only about 0·004 for LZ-4a while LZ-5a has a figure of 0·02.

Resistivity. The d.c. resistivity of the modified ceramic is much higher than the unmodified form. At 200° C the two resistivities differ by three orders of magnitude. Fig. 5 shows typical results for LZ-4a and LZ-5a over the temperature range of 25 to 300° C.

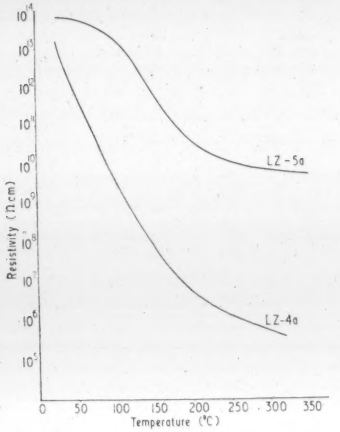


Fig. 5. Resistivity with temperature for ceramic types LZ-4a and LZ-5a.

Dielectric and elastic constants. The dielectric constant of the modified ceramic is higher, while its elastic stiffness is somewhat lower than that of the basic ceramic. This is reflected in the lower electrical and mechanical Q factors, the latter being about 75 for LZ-5a and 600 for LZ-4a.

Hysteresis. The modified material exhibits a relatively square hysteresis loop with coercive voltage which is quite independent of the applied field. Most ferroelectric ceramics show flattened hysteresis loops and usually the measured coercive field depends largely on the peak driving field. Generally the dielectric and mechanical properties of the ferroelectric crystals below their Curie points are functions of the state of polarization and stress.

A theory due to Devonshire (1954) has been quite successful in correlating many of the characteristics of the barium titanate single crystal, but the ceramic is more complicated. In a ceramic ferroelectric each crystallite is twinned into many polarization domains. The contribution of domain–wall motion to the properties is superimposed on the single crystal behaviour, and in many of the ceramic modifications this domain contribution is of considerable importance. In particular it is usually assumed that ageing is due to a gradual change in the domain configuration. Electrical hysteresis effects, whether in the single crystal or in the ceramic are also a characteristic due to domain wall motions.

Lead zirconate–titanate compositions with the described additives show more major differences in hysteresis and ageing than those found in the unmodified material. This may indicate variations in domain wall behaviour. Furthermore, most of the other variants in the modified material can be regarded as influenced by domain wall motion. For example, the high dielectric constants found in ferroelectric ceramics has been considered to be partly due to a limited domain wall oscillation, and in the case of 90° domain walls will result in increased compliance. Plessner (1956) has pointed out that a wall trapped in an unsymmetrical potential well will contribute to the power factor as well as to the dielectric and elastic susceptibilities. Thus, many of the unusual properties of the modified ceramic can be accounted for by the assumption that in this material the domain walls possess free movement under low alternating electrical or mechanical stress, but that the average position of the wall with no external stress applied is quite stable.

A series of experiments have been reported by Gerson (1960) to test these hypothesis. The approach used was to evaluate the ceramic microscopically and under electrical and mechanical excitation. This would tend to reveal domain behaviour. In these studies the ceramics used were of an unmodified form with the formula $Pb_{0\cdot95}Sr_{0\cdot05}Zr_{0\cdot55}Ti_{0\cdot47}O_3$ and a modified ceramic $PbZr_{0\cdot52}Ti_{0\cdot48}O_3$ + 1 wt. % Nb_2O_3. Both these materials exhibit a tetragonal form but may contain rhombohedral phases in minor amounts. The average dielectric constant for the unpolarized ceramic at room temperature is 1100 in the case of unmodified materials and 1200 for modified mixes. With polarized ceramics the free dielectric constants at 1 kc/s are 1250 parallel to the polarizing axes and 1350 perpendicular to this axis for the unmodified material while the corresponding values are 1500 and 1300 for the modified material. These values were taken 24 hours after polarizing. Figs 6 and 7 show the

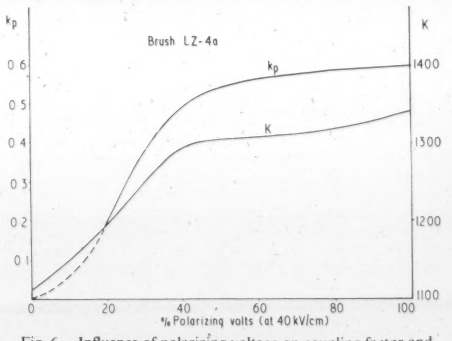


Fig. 6. Influence of polarizing voltage on coupling factor and dielectric constant for ceramic type LZ-4a.

effects of the degree of polarizing on the dielectric constant and planar coupling factor.

Figure 8 shows an electron micrograph of an unpolarized modified ceramic and both grain and domain boundaries are evident. The series of parallel domain boundaries are walls between domains where the polarization vectors are at 90° to each other, while the irregular lines within a grain are usually 180° walls. In spite of the difference in domain sizes, the average being 3 microns, the twinning patterns observed are similar to those found in barium titanate (Kulcsar 1956). For example, Fig. 9 shows two 180° walls crossing a field of 90° twinning. This pattern has been previously observed in barium titanate.

In his paper Gerson has proposed a hypothesis for the changes in modified lead zirconate–titanate based on the relatively high domain wall mobility in response to applied

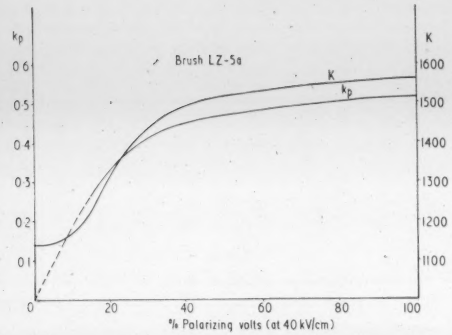


Fig. 7. Influence of polarizing voltage on coupling factor and dielectric constant for ceramic type LZ-5a.

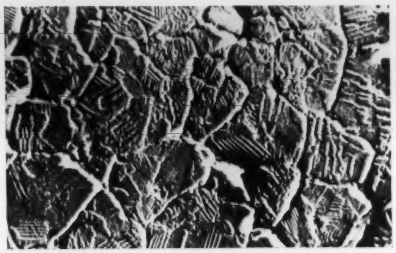


Fig. 8. Electron micrograph of unpoled modified ceramic.

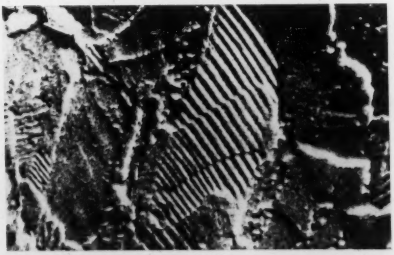


Fig. 9. Electron micrograph of unpoled modified ceramic showing 180° domain reversal in a field of 90° twinning.

electric fields. The common feature of the low ageing materials, both with three and five valent substituents is the deliberate introduction into the lattice of materials which do not possess the proper valence for the lattice position they probably assume. Based on the ionic radii of the elements it would be expected that the substitution sites of the additions would be lathanum instead of lead and niobium instead of titanium and zirconium. The effect of these substitutions is to cause vacancies which maintain the valency balance in the lead position. These vacancies, in a material whose ferroelectric properties are not destroyed by the disorder they introduce, may be expected to have a profound effect on domain wall motion under electrical or mechanical stress. This is because the elongation of the unit cell in the direction of the polarization results in very high local stresses on domains which have undergone 90° switching. These stresses may be relieved as switching continues, by changes in the overall dimensions of the ceramic. The presence of vacancies in the crystal minimizes the local stresses and enables switching to occur under low electrical or mechanical drive. The explanation of the reduced ageing of the modified ceramics as seen in Fig. 4 is that they are unable to store high stresses since these are quickly relieved by domain motion.

Domain wall motion is usually accompanied by hysteresis and is inherently a lossy process. Since there is a high mobility of the domain walls in modified material it is not surprising that under low electrical and mechanical stress this ceramic shows an increase in dielectric loss and a low mechanical Q.

Practical ceramics

While a large variety of characteristics are possible with different modifications of the basic lead zirconate–titanate it has been found that initially three types can be standardized as possessing the most advantages and the least disadvantages for general applications. The comparative properties are summarized in the Table.

Low signal parameters of piezoelectric ceramic compositions

		$BaTiO_3$	LZ-4a	LZ-5a	LZ-6
Coupling coefficients					
	k_{33}	0·50	0·76	0·68	0·54
	k_p	0·36	0·55	0·54	0·39
	k_{31}	0·21	0·30	0·32	0·229
Piezoelectric strain constants (10^{-12}C/N)					
	d_{33}	190	300	320	191
	d_{31}	−78	−130	−140	−78
Free dielectric constants					
	K_3	1700	1200	1500	1075
	dissipation factor	0·01	0·005	0·02	0·02
	K_1	1450		1285	
Electric constants (10^{10}N/m^2)					
	$1/S_{33}^E$	10·5	6·75	5·9	—
	C_{33}^D	17·6		13·5	
	$1/S_{11}^E$	11·0	6·85	6·8	8·7
Density					
	ρ	5·7	7·6	7·5	7·5
Propagation velocity (m/s)					
axial	C_{33}^D/ρ	5600	4500	4200	—
lateral	$1/S_{11}^E.\rho$	4400	3040	3000	3400
shear	C_{55}^E/ρ	2800	1860	1650	—
Mechanical Q (planar)					
	Q_m	400	500	75	340
Resistivity at 100° C (MΩ m)		30	300	10^5	400

Type LZ-4a has a moderately high dielectric constant and resembles barium titanate in its general uses, although superior for all of them. It is best suited for use as a driving transducer or resonator with a relatively high mechanical Q and efficient electromechanical coupling. The ageing effect in LZ-4a is very similar to that of barium titanate and so is the electrical resistivity at elevated temperatures. Some piezoelectric and electrical characteristics are shown in Figs 10 and 11 as a function of temperature.

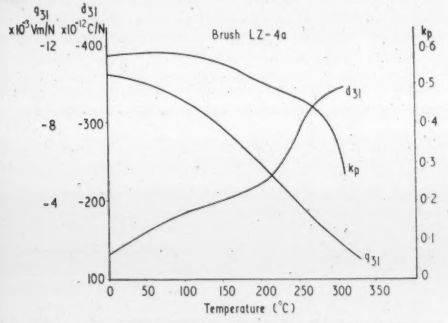


Fig. 10. Type LZ-4a ceramic piezoelectric characteristics as a function of temperature.

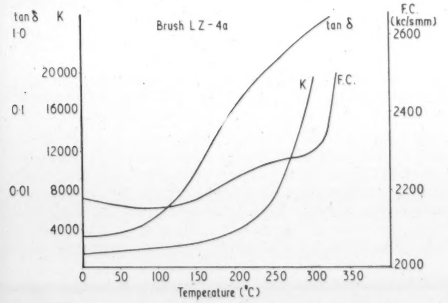


Fig. 11. Type LZ-4a ceramic electrical characteristics as a function of temperature.

For high power operation of piezoelectric transducers the dielectric losses must be low. With a low loaded mechanical Q it is obvious that it is necessary to use high driving fields to achieve a high stress. Dielectric losses produce heat energy and thus reduce acoustic efficiency. If this heat is not conducted away from the transducer it may cause a temperature rise sufficient to destroy polarization in the ceramic. In most cases the dielectric loss at a given field increases with the temperature and the effect rapidly worsens. With LZ-4a it is possible to drive at a strain amplitude about twice as great as with most barium titanate compositions. The high Curie point enables it to be operated safely at levels far above other materials and this also contributes to the power handling abilities.

Figure 12 shows the relation between heat power dissipation expressed as internal heating per cubic metre per cycle per second and mechanical strain under non-resonant driving conditions. It will be seen that at 200° C the LZ-5a ceramic is superior to the unmodified material for low strain amplitudes. Here the heat losses due to d.c. conductivity make a

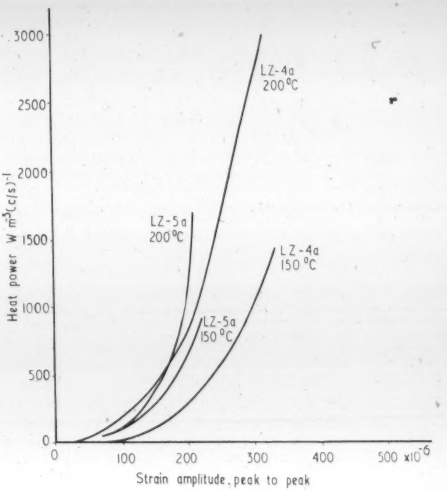


Fig. 12. Internal heating as a function of mechanical strain for LZ ceramics.

major contribution with LZ-4a. The modified ceramic possesses a much lower conductivity at 200° C and is therefore not subject to these losses.

Type LZ-5a has an even higher coupling coefficient than the unmodified material and also possesses a high dielectric constant. It is ideally suited for use in gramophone pick-ups, accelerometer transducers and other non-resonant conditions of operation. The mechanical Q is low in comparison with type LZ-4a and thus precludes its use as a driver. Increasing dielectric losses at high driving fields develop internal heat and the room temperature dielectric loss angle is higher than LZ-4a. The ageing effect in this composition is small compared with other ceramics. Figs 13 and 14 show the piezo-

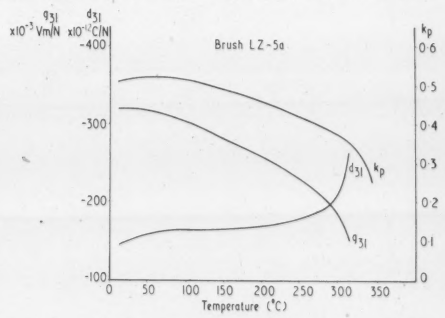


Fig. 13. Type LZ-5a ceramic piezoelectric characteristics as a function of temperature.

electric and electrical characteristics of this material over a wide temperature range.

Finally, a type LZ-6 has been developed specifically for maximum stability of resonant frequency and is finding increasing use as a frequency filter material. The resonant frequency is almost independent of time and temperature

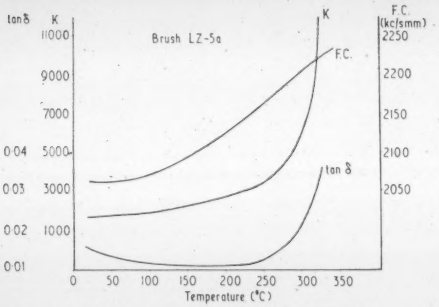


Fig. 14. Type LZ-5a ceramic electrical characteristics as a function of temperature.

variations. The dielectric constant and coupling coefficients are not as great as those of the type LZ-4a and 5a mixes but nevertheless the coupling coefficients over the range −60 to 250° c is higher than that of commercially available barium titanate at room temperature. As with the type LZ-5a material LZ-6 has very low ageing but possesses the high Q characteristics of LZ-4a. The resonant frequency of stabilized LZ-6 is constant within $\pm 0{\cdot}15\%$ from −60 to 200° c and is virtually unchanged by shelf storage after one week from polarizing.

Conclusions

The improved characteristics of the lead zirconate–titanate ceramics are enabling many new applications to be exploited which were not previously possible with barium titanate. The very high transverse coupling factor opens up possibilities in the design of driving transducers where the relationship of the transverse mode to the longitudinal mode can be fully utilized. For example, very efficient drivers can be constructed using tubes energized across the wall thickness but operating in the length mode. Similarly, the use of thickness polarized but length energized bars has recently been applied in the transverse type piezoelectric voltage transformers (Crawford 1960).

Apart from acoustic devices certain characteristics of LZ ceramics have been exploited in other fields of application. A recent development uses the high remanent polarization for the generation of high voltage by the conversion of mechanical force, and practical applications have been found as an ignition system in internal combustion engines. Until the recent work of Jaffe, Gerson and their co-workers it was difficult to understand the relatively complicated processes involved in modified ceramics. It has already been possible to forecast performance alterations due to additives and this will give increasing opportunities in the manufacture of a piezoelectric ceramic to a desired performance specification and thus rapidly widen the scope of application.

Acknowledgments

Most of the investigation work outlined in this paper has been carried out at the Research Laboratories of the Clevite Corporation, under the direction of Dr. Hans Jaffe. Sincere appreciation is given to Dr. Jaffe for his continued encouragement of work undertaken at Brush Crystal Co. Ltd. The writer must also acknowledge the efforts of his own colleagues, notably Dr. J. A. Sugden in charge of Ceramics preparations and Mr. R. F. J. Orwell who carried out many of the measurements.

Thanks are also given to the Directors of Brush Crystal Co. Ltd., for permission to publish this paper.

References

CRAWFORD, A. E., 1959, *Brit. Comm. Elec.*, **6,** 516.
—— 1960, *Wireless World*, **66,** 510.
DEVONSHIRE, A. F., 1954, *Phil. Mag. Suppl.*, **3,** 85.
GERSON, R., 1960, *J. Appl. Phys.*, **31,** 188.
GRAY, R. B., 1949, U.S. Patent 2, 486, 560.
JAFFE, B., ROTH, R. S., and MARZULLO, S., 1954, *J. Appl. Phys.*, **25,** 809.
KULCSAR, F., 1956, *J. Amer. Ceram. Soc.*, **39,** 13.
—— 1959a, *J. Amer. Ceram. Soc.*, **42,** 49.
—— 1959b, *J. Amer. Ceram. Soc.*, **42,** 343.
PLESSNER, K. W., 1956, *Proc. Phys. Soc.*, B, **691,** 1261.
SHIRANE, G., and SUZUKI, K., 1952, *J. Phys. Soc. Japan*, **7,** 333.
WAINER, E., 1946, *Trans Electrochem. Soc.*, **89,** 331.

Magnetism in transition metal alloys

by W. M. LOMER, M.A., M.Sc., Ph.D., Atomic Energy Research Establishment, Harwell, Didcot, Berks.

MS. received 29th June 1961

Abstract

The electronic structure of alloys involving transition elements is complex, but recent studies by various techniques allow us to discuss the range of validity of certain simple concepts. This paper emphasizes the domain of validity of the idea of a local atomic spin and points out that the collapse of this model is associated with the lifetimes of the spins. When the lifetime becomes very short the model is entirely inappropriate. It appears, however, to apply to the metals Cr, Mn, Fe, Co, Ni and to their alloys with many simple metals.

THE transition elements in the first long period show strong magnetic properties, which are quite unlike those of typical metals such as the sodium, magnesium, aluminium, copper, zinc and gallium groups. The ferromagnetic trio Fe, Co, Ni have, of course, long been recognized as unusual, and since the advent of nuclear reactors to provide neutron beams, it has been shown that α-Mn (Shull and Wilkinson 1953) and Cr (Bacon 1961) are antiferromagnetic at low temperatures. Atoms in these structures have magnetic moments associated with them which are oppositely directed so that no net magnetization is recognizable. In the ferromagnetic elements it is possible to determine the ratio of magnetic moment to mechanical moment of inertia, either by direct means in the Einstein–de Haas experiment, or by the far more accurate ferromagnetic resonance technique (Kittel 1956a). The result indicates that nearly all of the magnetism arises from electron spin, and scarcely any contribution comes from orbital motion.

Theoretical considerations and comparison with the properties of paramagnetic salts of these elements suggest strongly that the electron group responsible is the incomplete 3d shell. Measurements of the atomic form factor for the magnetic neutron scattering from manganese compounds supports this with some accuracy (Hastings *et al.* 1959), and there is no reason to doubt that strong magnetism is entirely associated with the spin of unpaired 3d electrons. The orbital motion of the 3d electrons is suppressed by the non-spherical nature of the environment of the atoms (Kittel 1956b), or by the band motion of the electrons (Brooks 1940).

The magnetic behaviour of a material containing a high density of interacting elementary magnets (electron spins) is well understood in qualitative terms. At high temperatures the spins are randomly oriented in zero field, and when a weak magnetic field is applied they are partially aligned to give a magnetic moment

$$M = \frac{\mu^2}{3k(T - \Theta)}$$

where μ = magnetic moment of elementary magnet

Θ = paramagnetic Curie temperature

$\simeq k^{-1} \times$ interaction energy between a spin and its fully aligned neighbours.

This is the familiar Curie–Weiss law, where a positive Θ corresponds to a tendency to ferromagnetism and a negative Θ to antiferromagnetism. In each case the law breaks down at $T \lesssim |\Theta|$, and in this region strong magnetic properties arise and may be dealt with roughly in the molecular field approximation and, at very low temperatures, in the spin wave approximation (Kittel 1956c).

Whilst Fe, Co and Ni give reasonable Curie–Weiss behaviour at high temperatures, the same is not true of Mn and Cr, both of which give much smaller temperature dependence of susceptibility than would be expected on the model of nearly independent spins. The value of the susceptibility in each case is also lower than such a model would suggest.

The low temperature susceptibility of a polycrystalline antiferromagnet consisting of elementary magnets of moment μ with total interaction of the order of $k\Theta$ per moment is given approximately by

$$\chi \sim \frac{1}{3}\frac{\mu^2}{3k\Theta}.$$

In non-transition metals the electron spins cannot be treated as independent elementary magnets because their wave functions overlap in space and so note must be taken of the Pauli principle, giving rise to the familiar degenerate electron gas properties. Under these conditions the number of electrons which have spins parallel to a given direction can be increased only at the cost of increasing their kinetic energy. The net susceptibility is much less than it is in the Curie–Weiss regime where the entropy was the only term opposing magnetization of the ensemble. The expression usually derived (without exchange correction) for the electron gas is

$$\chi = \mu^2 N(E_0)$$

where $N(E_0)$ is the density of electron energy states at the Fermi energy E_0. Since $N(E_0)$ is roughly $1/E_0$ and E_0 is a few electron volts, this is very much less than a Curie–Weiss value.

The early transition elements Sc, Ti, V, all show susceptibilities nearly independent of temperature, and no evidence of strong magnetism, so that we may conclude that in these elements at least the unfilled d electrons are treated in reasonable approximation by band theory, and the same may be true at high temperatures for Cr and Mn.

In view of the situation outlined above, it is of very considerable interest to study the effect of diluting the apparently spin-bearing atoms Cr, Mn, Fe, Co, Ni with other metals to determine the conditions under which a local atomic moment may exist. The alloys fall into four classes.

(i) Alloys based on non-magnetic solvents with low density of states (e.g. Cu based).
(ii) Alloys based on high density of states solvents (e.g. Al, V).
(iii) Alloys with Cr (few elements dissolved in α-Mn).
(iv) Alloys with ferromagnetics.

Typical of group (i) are the alloys **Cu** Mn and **Au** Mn. They show a magnetic susceptibility following a Curie–Weiss law temperature dependence down to a low temperature. Then they become antiferromagnetic despite the randomness of the arrangement of the moment-bearing atoms (Owen *et al.* 1957). Their low-temperature thermal and magnetic properties are apparently quite in accord with the model of a number of well-defined weakly interacting spins (Marshall 1960). The value of the Néel temperature for these alloys is dependent on concentration, and for a few per cent Mn in **Cu** is about 10° K. This corresponds to mean interactions of about 10^{-3} ev between spins, so that at temperatures above the spin ordering temperature spins 'exchange' or flip one another with frequencies as high as $h^{-1}(k\,.\,10) \sim 10^{12}$ s^{-1}. Thus the spins will not be properly identified in unmagnetized samples by experiments which observe effects averaged over considerable periods. It is for this reason that ^{57}Fe dissolved in copper, or indeed any other non-magnetic metal such as stainless steel, affords a Mössbauer soft γ-ray emission source in which the iron nucleus experiences no effective magnetic field (Wertheim 1960). The emission process requires 10^{-7} second for completion, during which time the iron atomic spin will have been flipped many times and will consequently have an average of zero. The flipping process in very dilute solutions probably involves the conduction electrons as well as other solute atoms.

It is convenient to follow the group (i) alloys with those based on Cr, which also has a low density of states, though it is at low temperatures antiferromagnetic. It has been shown (Lomer 1960) that the magnetic properties of Fe and Co dissolved in Cr can be explained if these atoms carry effective spins which interact only rather weakly with the antiferromagnetic lattice of the matrix. It seems that conduction electron interaction is insufficient to destroy the thermodynamic significance of the solute spin state. One would expect that until the temperature lay below that corresponding to such interaction the spins would again undergo spin flips very rapidly on the time scale of Mössbauer emission or nuclear magnetic resonance.

Group (ii) above—dilute alloys based on a matrix which has a high density of states—show magnetic properties typical of a degenerate electron gas. The susceptibility is temperature independent and its variation appears to be dependent on the total number of valence electrons added (Gardner and Childs, to be published). The interpretation of this in general terms may be that any spin on a dissolved atom interacts so strongly with the conduction band as a whole that its lifetime is less than h/kT—the spin levels would appear to be broader, because of their short lifetime, than the Fermi surface thickness. In fact, it is easy to see that this is not an admissible way of speaking. The spin and band states must be mixed to give a complicated set of states which interact more weakly than kT. Though it seems very plausible then that the whole ensemble will behave like an electron gas, the exact form for the new density of states in the alloy is not known.

The ferromagnetic alloys form a subject too complex to be discussed here, but a summary of their magnetic properties and an empirical classification scheme was put forward by Marshall and the author (Lomer and Marshall 1958). While several features of that paper are now known to be inadequate, the general discussion appears to be still valid in broad outline.

In several alloys which show apparently degenerate electron gas behaviour at low concentrations, higher concentrations show temperature dependent susceptibilities. This is observed, for example in **V** Fe and **V** Mn (Gardner, to be published). Cheng, Wei and Beck (1960) have shown that at similar composition the low temperature linear specific heat term reaches anomalously high values. This happens in **V** Fe at concentrations just below those needed for the first appearance of ferromagnetism. It seems likely that in some way the presence of other atoms with spin bearing potentialities stabilizes the spin states and weakens their interaction with the electron gas to such an extent that spin groupings do acquire a lifetime great enough to justify the random effective field treatment used by Marshall to explain the high linear specific heat term in copper–manganese alloys (Marshall 1960).

The conclusion of this discussion is to be summarized in a few generalizations, which further work will confirm, or modify, or destroy. Let us simply list them:

(i) The atoms from Cr to Ni carry spin moments in suitable environments.

(ii) The lifetime of such spin states is controlled by interactions with other spins, and with the conduction band. At low temperatures interactions with other spins give rise to very long lifetime, strong magnetic states, ferro- or antiferromagnetic; at high temperatures these interactions, and at *all* temperatures conduction electron interactions, shorten the lifetimes.

(iii) The main factor controlling the conduction electron interaction is the density of states.

(iv) For a thermodynamic treatment of the spins to be valid, their lifetime must exceed h/kT, but much more stringent conditions are required in experiments involving longer characteristic times.

It will be interesting to watch the development of these ideas and the evolution of a rapprochement between them and those of the Friedel school (Friedel 1954, 1956, de Gennes and Friedel 1958) in the course of the next few years.

References

BACON, G. E., 1961, *Acta Cryst.*, in the press.

BROOKS, H., 1940, *Phys. Rev.*, **58**, 909.

CHENG, C. H., WEI, C. T., and BECK, P. A., 1960, *Phys. Rev.*, **120**, 426.

FRIEDEL, J., 1954, *Advanc. in Phys.*, **3**, 446.

—— 1956, *Canad. J. Phys.*, **34**, 1190.

DE GENNES, P. G., and FRIEDEL, J., 1958, *J. Phys. Chem. Solids*, **4**, 71.

HASTINGS, J. M., ELLIOTT, N., and CORLISS, L. M., 1959, *Phys. Rev.*, **115**, 13.

KITTEL, C., 1956a, *Solid State Physics* (New York: Wiley), p. 410.

—— 1956b, *Solid State Physics* (New York: Wiley), p. 219.

—— 1956c, *Solid State Physics* (New York: Wiley), p. 405 *et seq.*

LOMER, W. M., 1960, *Australian J. Phys.*, **13**, 451.

LOMER, W. M., and MARSHALL, W., 1958, *Phil. Mag.*, **3**, 185.

MARSHALL, W., 1960, *Phys. Rev.*, **118**, 1519.

OWEN, J., BROWNE, M., ARP, V., and KIP, A. F., 1957, *J. Phys. Chem. Solids*, **2**, 85

SHULL, C. G., and WILKINSON, M. K., 1953, *Rev. Mod. Phys.* **25**, 100.

WERTHEIM, G. K., 1960, *Phys. Rev. Letters*, **4**, 403.

The reactor as a source of radiation

by A. J. SALMON, B.Sc., Ph.D., A.M.I.E.E., F.Inst.P., Research Laboratory, Associated Electrical Industries Ltd., Aldermaston Court, Aldermaston, Berks.

MS. received 24*th March* 1961, *in revised form* 28*th June* 1961

Abstract

The paper surveys the experimental facilities and the intensities of the radiations which are now available in research reactors and discusses some of the characteristics of the various reactor types of interest to the radiation worker. The types and numbers of general purpose research reactors now available throughout the world are also indicated, and some particular applications of reactors for research are described.

1. Introduction

THE nuclear reactor is one of the most intense sources of neutrons and γ-radiation available; as an intense source of neutrons of all energies it is unequalled. An accelerator is superior when neutrons of one specific high energy, or a very short pulse of neutrons, is required, and the linear accelerator is a comparable source of γ-radiation. However, the reactor can provide γ-radiation and neutrons at the same time for many different experiments, via the many experimental facilities which can be arranged around a reactor (see Fig. 1). Indeed one of the difficulties associated with many reactor experiments lies in reducing the intensity of undesired radiation.

The reactor is increasing in importance as a research tool and this article surveys the general considerations pertinent to this field of inquiry. More detailed considerations, of particular points, can be found in the appropriate textbooks, for example: Hughes (1953), Glasstone and Edlund (1952), Chastain (1958) and Glasstone (1956).

2. Types and distributions of research reactors

Several parameters characterize a particular reactor. The most important are: the neutron spectrum, the moderator,

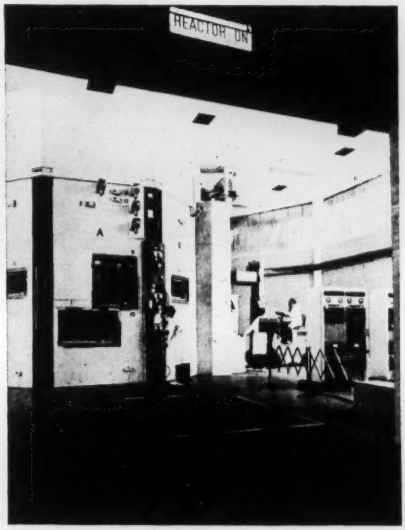


Fig. 1. A typical medium sized research reactor showing several experimental facilities. In the centre are two 6 inch and one 12 inch diameter experimental tubes, then slightly to the right another 6 inch diameter experimental tube, whilst on the far right there is a thermal column.

Table 1. *Major types of research reactors*

(1)	(2)	(3)	(4)	(5)	(6)
1	Natural U, graphite	0·7	0·1–30	large	BEPO, GLEEP
2	Enriched U, graphite	~90	10–30	large	Brookhaven (U.S.A.)
3	Near nat. U, D_2O	0·7 to 1·5	10–200	large	NRU (Canada)
4	Enriched U, D_2O	20 to ~80	1–10	in tank	DIDO, M.I.T. (U.S.A.)
5	Enriched U, H_2O	20 to ~90	0·1–5	in open pool	MERLIN, HERALD
6	Enriched U, H_2O	2 to ~90	10–200	in tank	MTR (U.S.A.)

(1) Type; (2) fuel, moderator; (3) $\%^{235}U$; (4) usual power range (MW); (5) physical characteristics; (6) examples.

if any, of the neutrons, the thermal power output, and the percentage of ^{235}U in the uranium fuel.

Various combinations of these parameters are used in different reactors. Those of the major research reactors are given in Table 1. In all of them the core neutron energy spectrum has a peak in the thermal region and all use uranium as fuel. In one class the main differences are the lower

Table 2. *Types of low-power training and research reactors*

All have thermal power outputs of 100 kw or less. All use enriched uranium

(1) *	(2)	(3)	(4)	(5)	(6)
7	Heterogeneous	H_2O	10–100	in pool	CONSORT
8	Heterogeneous	ZrH	10–100	in pool	TRIGA
9	Heterogeneous	H_2O/graphite	0·1–100	in tank	JASON
10	Homogeneous	H_2O	10^{-2}–50	small	JRRI (Japan)
11	Homogeneous	Polystyrene or graphite	10^{-3}–10^{-2}	small	AGN 211

(1) Type; (2) fuel; (3) moderator; (4) usual power range (kw); (5) physical characteristics; (6) examples.

output, to which the radiation intensities are approximately proportional, and the experimental facilities.

Another group of reactors, of relatively low power and designed as much for training as for research, is listed in Table 2. Their low neutron flux limits their application in research pertinent to nuclear power, but they should be very useful in the field of general radiation research.

The main individual characteristic of any research reactor is its power. The power levels of the world's research reactors are shown in Table 3.

Table 3. *The distribution of research reactors throughout the world*

The reactors were in operation, under construction or definitely planned in May 1961

Power (w)	*U.K.*	*Europe**	*U.S.A.*	*Rest*	*Total*
$\leqslant 1$	0	2	12	0	14
$>1, \leqslant 10$	0	1	9	1	11
$>10, \leqslant 10^3$	0	3	9	0	12
$>10^3, \leqslant 10^5$	6	14	38	15	73
$>10^5, \leqslant 10^6$	0	11	13	8	32
$>10^6, \leqslant 5\times10^6$	2	24	12	5	43
$>5\times10^6, \leqslant 10^7$	4	6	3	5	18
$>10^7, \leqslant 5\times10^7$	0	10	6	3	19
$>5\times10^7$	0	0	5	1	6

* including U.S.S.R.

3. Types of radiation

The neutron velocity spectrum in the core of a research reactor will vary from point to point, but in general it will be similar to that shown in Fig. 2. The main features are:

(*a*) The near Maxwellian distribution, with a peak at a velocity corresponding approximately to the temperature of the moderator. The neutrons composing this distribution are usually called 'thermal' neutrons. Those in the low energy end of this group are sometimes called 'cold' neutrons as their energies correspond to temperatures less than that of the bulk material. The mode of the thermal neutron velocity distribution when the temperature is 20·44° c will have a corresponding energy of 0·0253 ev, equivalent to kT, and a neutron velocity of 2200 m/s; cross sections for thermal neutrons are often tabulated for this particular velocity.

(*b*) Merging into the Maxwellian distribution and extending up to the highest energies are the 'epithermal' neutrons. These epithermal neutrons are those neutrons which are not part of the Maxwellian distribution. Theory, neglecting resonances and variations in the scattering cross section with energy, predicts an epithermal energy spectrum of the form

$$n(E)dE \propto dE/E^{3/2},$$

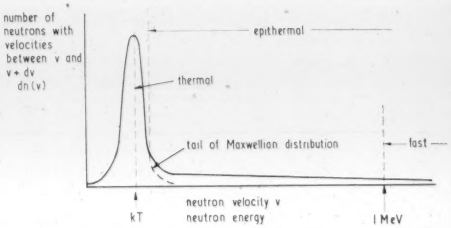


Fig. 2. A typical neutron spectrum in a reactor. Note that there is no particular energy limit distinguishing thermal neutrons from epithermal neutrons.

where $n(E)dE$ is the number of neutrons per unit volume in the energy interval dE, for neutron energies above a few kT. The corresponding velocity spectrum is

$$n(v)dv \propto dv/v^2.$$

This epithermal neutron spectrum will extend upwards in the energy scale until at energies of above 1 MeV it will be practically indistinguishable from the spectrum of neutrons from the fission process.

(*c*) The epithermal neutrons with energies above 1 MeV, probably uncollided fission neutrons, are called 'fast' neutrons, indeed, this is better stated as 'fast (>1 MeV) neutrons'. There is no fully accepted convention yet about the above terms, especially about what is meant by epithermal and fast neutrons; however the above are acceptable if their meaning is stated.

Beta-rays and electrons are abundantly available for irradiation experiments; if a particular β-ray is required then the pertinent isotope probably can be produced in the reactor.

Čerenkov radiation is available in research reactors having a liquid as moderator or coolant. In a water cooled and moderated reactor operating at 1 MW the amount of energy appearing as Čerenkov radiation will amount to about 5 W. The radiation is available in the reactor tank or elsewhere in the primary coolant circuit. Although not a good source for investigations of Čerenkov radiation itself, it is a good source for some experiments involving the detection of the gross effect.

4. Units

The measure of the intensity of neutrons is the neutron flux. Considering neutrons of one velocity only, v cm/s, their density being n/cm^3, then the flux is defined as nv/cm^{2}s. This is the number of neutrons passing through a sphere, whose cross-sectional area is one square centimetre, per second. Expressed in other terms the flux is equal to the track length traced out by the neutrons in one cubic centimetre per second. The true total flux will be given by

$$F = \int_0^\infty f(v)dv,$$

where the flux distribution function $f(v) = vn(v)$. In practice the fluxes usually quoted are slightly different from this true flux, but the difference is small; the topic has been discussed by Westcott, Walker and Alexander (1958).

Gamma-radiation intensities in and about reactors are usually given in rads (tissue) per hour, or in röntgens per hour. γ-ray doses are quoted in rads. However, the neutron irradiation received by a specimen is usually quoted in neutrons/cm^2. This is the neutron flux in which the specimen has been, multiplied by the time in seconds during which it has been irradiated, e.g. if a specimen has been in a fast (>1 MeV) neutron flux of 10^{14}/cm^{2}s for 10 days ($\sim 10^6$s) then the fast neutron dose would be stated as about 10^{20} neutrons/cm^2. Sometimes this dose would be stated as 10^{20} nvt, indicating that it is a product of the flux and time.

5. Typical facilities and radiation intensities

The main facilities, available in almost all research reactors, are described in the following sections. Fig. 3 shows a

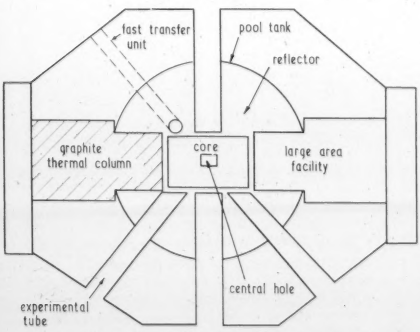


Fig. 3. A typical plan of a pool type reactor showing the various types of experimental facility.

typical arrangement in a pool type reactor. The reactor core, consisting of the fuel, moderator and coolant, is at the centre of the pool. The core is surrounded by the reflector which is often of the same material as the moderator in the core. The other facilities penetrate the shield, about the core and reflector, ending as near the core as possible.

5.1. *The reactor core*

Here the radiation intensities are highest and, often, experimental space is severely limited. Only in reactors using natural, or near natural, uranium as fuel is there plenty of space in the core.

The undisturbed neutron fluxes (thermal, epithermal and fast) in the core follow, approximately, a cosine distribution in both the vertical and horizontal directions. However, there will be fine structure effects in a heterogeneous reactor due to the different properties of the fuel, the moderator and the coolant. Fig. 4 (K. Stephens and G. Williams 1960,

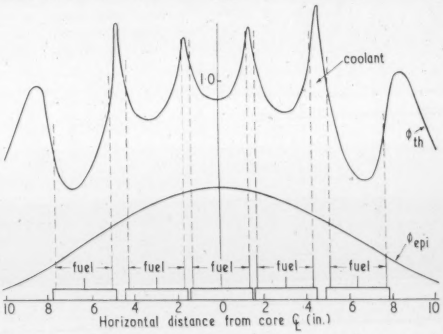


Fig. 4. Relative horizontal thermal and epithermal neutron flux distributions in the midplane of MERLIN.
1·0 unit = 1·33 × 10^7 thermal neutrons/cm^{2}s per watt.
1·0 unit = 2·95 × 10^6 epithermal neutrons/cm^{2}s per watt.

unpublished), shows this fine structure of the thermal flux in a water moderated and cooled reactor, using highly enriched fuel, the peaking of the distribution in the coolant channels and the depression in the fuel are marked. The epithermal and fast fluxes are not subject to this fine structure to the same extent.

For a given type of reactor the radiation intensities are approximately proportional to the thermal power output per unit volume. The volume of the core is, mainly, a function of the moderator and the fuel enrichment; with a low ^{235}U percentage content the volume is increased. Of the three most popular moderators, graphite usually provides the most space, then heavy water; light water produces the most compact cores (see Fig. 5). However, light water has the

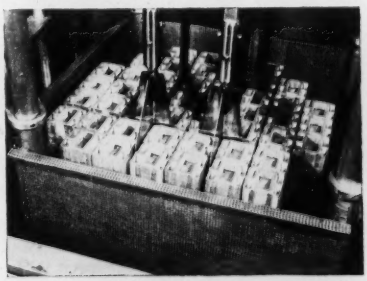


Fig. 5. The core of a water moderated and cooled reactor is very compact. The core shown contains 2·0 kg ^{235}U and is only about 2 ft × 1 ft × 1 ft; in this particular case it is surrounded by a beryllia reflector. Four control elements can be seen in the core.

advantages of simplicity and cheapness and also results in a higher maximum fast flux per unit of reactor power; the figures for DIDO and MERLIN are $0{\cdot}25 \times 10^{13}$/cm^{2}s MW and $1{\cdot}5 \times 10^{13}$/cm^{2}s MW respectively. The fast flux in DIDO has been considerably enhanced by placing specimens inside hollow fuel elements. By this means fast fluxes of about 2×10^{13}/cm^{2}s MW have been obtained for the irradiation of specimens (B. O. Wade 1960, unpublished).

The γ-radiation absorption in the core of a reactor which has been operating for a considerable time, assuming no leakage out, is about 4×10^{11} MeV/cm^{3}s per w/cm^3. In a light water cooled and moderated reactor using highly enriched fuel, at 5 MW, this corresponds to a dose rate of 4×10^8 r/h and an energy input of about 0·7 w/g for small specimens near the centre of the core. Such high rates of γ-ray heating, in cores with high power densities, produce severe problems in the design of experimental equipment.

Another problem, concerning all reactor experiments, is that of shielding. Specimens or apparatus brought out of the neutron fluxes have to be shielded so that they do not damage personnel or disturb other experiments. Similarly, if radiation beams are allowed into the experimental area, shielding has to be provided.

No rules can be given for the radiation intensities available, not even for one type of research reactor; as we have seen they may be increased markedly by various means, they do vary markedly from experimental facility to facility and, indeed, they may be lower than expected because of the design of the reactor. However, Table 4 indicates the maximum intensities that might be expected in the core of a particular type of reactor. Some reactors will be better, some will be worse than indicated here.

generation to the next. Thus it is closely associated with the safety of the system.) Also the γ-ray heating of objects in the reflector will be less. Thus, experiments are easier to design and safer than if they were in the core. However the maximum fast flux, the maximum ratio of fast to thermal flux and the minimum gradient of a radiation intensity are in the core.

It is often desirable to optimize the intensity of one type of radiation relative to the other types, for example, a flux as near thermal as possible may be required with as little as possible of γ-radiation or fast neutrons present. Some variation of the ratios of the radiation intensities is possible by the choice of experimental facility. To obtain a larger ratio than exists in the reactor normally, special techniques can be used employing combinations of filters and converters, see Fig. 6. To optimize the thermal neutrons a 'thermal

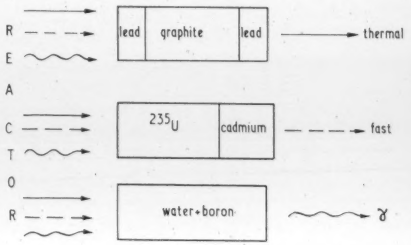


Fig. 6. Idealized schemes for producing 'clean' radiation from the mixture of radiations existing in a reactor.

Table 4. *The maximum radiation intensities, per unit power, available in a reactor core*

The types are those given in Tables 1 and 2

Type	1	2	3	4	5	6	7	8	9	10	11
Thermal, 10^{12}/cm^{2}s MW	0·2	1	2	12	10	7	15	20	15	20	35
Fast (>1 MeV), 10^{12}/cm^{2}s MW	0·02	0·04	0·1	2	10	5	15	10	15	20	30
γ, 10^7 rad/h MW	0·01	0·02	0·1	2	8	4	12	8	12	15	25

5.2. *The reflector*

The reflector is the region surrounding the core, its minimum thickness may vary from between 2½ in. in a light water reflected reactor to about 3 ft in a heavy water reflected reactor, the maximum thickness may be many feet especially in a reactor of the swimming pool type. There is usually an increase in thermal neutron flux near the core–reflector boundary (Fig. 4). After this peak the thermal flux will decrease approximately exponentially. The epithermal neutron, the fast neutron, and the γ-radiation intensities all decrease continuously with distance into the reflector.

Depending on the arrangement of facilities about the core, it may be possible to put large experiments into the reflector. In pool type reactors one face of the core may be used for placing experiments directly against that face. In many cases the thermal flux at the interface is as high as at the centre of the core and this makes the reflector space valuable as it is usually easier to place the experiment in the reflector. Another important advantage of the reflector is that experiments in this region will have less effect on reactivity than if the same experiments were placed in the core. (When a reactor is near to the critical state, the reactivity is the fractional change in the number of neutrons from one column' of graphite is used: there is one in almost every research reactor, see below. To optimize the fast neutrons a converter cylinder or plate of ^{235}U may be used, the fission of the ^{235}U by the thermal neutrons producing the fast neutrons. A hollow fuel element of highly enriched uranium is an example of such a converter, see Fig. 7. Cadmium can be used to absorb any thermal neutrons which have

Fig. 7. A fuel element of highly enriched uranium in which small samples can be irradiated in a high fast flux. This type of fuel element is used by the U.K.A.E.A. in their heavy water reactors. Specimens, of diameters up to two inches, can be accommodated inside the element.

by-passed the ^{235}U. Water is a useful medium for removing the thermal and fast neutrons, whilst leaving some γ-radiation available; the addition of a boron compound attenuates the thermal neutron intensity even more rapidly.

5.3. *Experimental tubes*

Experimental tubes extend from outside the reactor shield to near the core. The diameter of the tube is normally between 1 in. and 12 in.; a few are larger; most experimental tubes in modern reactors have diameters between 4 in. and 8 in. They can be used either to permit a beam of radiation to escape into the experimental area around the reactor or as an irradiation facility. In the latter case the experiment is usually placed within the tube as near the core as possible. Neutron flux distributions down an experimental tube are shown in Fig. 8, γ-ray distributions are shown in Fig. 9. The fast neutron spectrum at the end of a tube near the core, in a water moderated reactor, is very similar to a fission

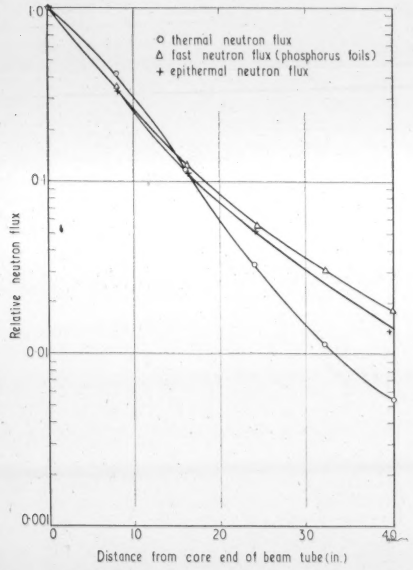


Fig. 8. Flux distribution in 6 in. diameter experimental tube. The end of the tube was separated from the core surface by about 7 in. of water (K. Stephens and G. Williams 1960, unpublished).

1·0 unit $\equiv 2{\cdot}3 \times 10^5$ thermal neutrons/cm^{2}s per watt.
1·0 unit $\equiv 6{\cdot}2 \times 10^4$ fast (>1 MeV) neutrons/cm^{2}s per watt.
1·0 unit $\equiv 7 \times 10^3$ epithermal neutrons/cm^{2}s per watt.

spectrum. This is sometimes moderated to enhance the 'cold' neutrons by placing refrigerated material, e.g. liquid hydrogen, in the tube. Similarly, it should be possible to enhance the low energy epithermal neutrons by placing hot material in the tube, e.g. graphite at 1000° C or more.

By arranging the tubes tangentially to the core rather than radially, it is possible to increase the ratio of the thermal neutron intensity to the intensity of fast neutrons and γ-rays in the beam; this is achieved only with some loss in the intensity of the thermal neutron beam. However, the loss might be worth while; in an experiment with a reactor reflected by heavy water, for a reduction in thermal neutron intensity of about 2 the fast neutron and γ-rays were reduced by a factor of about 15 (Court and Downes 1960).

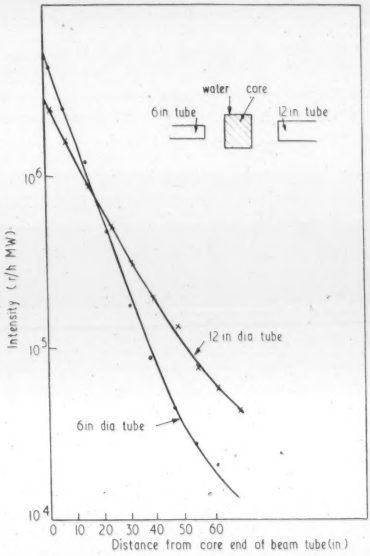


Fig. 9. Gamma-radiation intensity down 6 in. and 12 in. diameter experimental tubes. The arrangement of the tubes with respect to the core is shown (J. Mitchell 1960, unpublished).

5.4. *Thermal columns*

Thermal columns are large facilities penetrating from the experimental area to the region of the core and filled with graphite, see Fig. 10. The fast neutrons are moderated by the graphite, which does not absorb thermal neutrons to any marked extent; the γ-radiation is absorbed by a lead sheet or, at a greater financial cost but more efficiently, by a bismuth sheet. The outside face of the thermal column can be used as a plane source, of large area, of thermal neutrons. These thermal neutrons will be relatively uncontaminated with fast neutrons and the γ-intensity can be made small. Typical dimensions for a thermal column are 6 ft long by 5 ft square. Over the length of this column the thermal neutron flux will have decreased by a factor of about 10^3, whilst the fast neutrons and γ-radiation will have decreased by factors of about 10^5 and 10^3 respectively. Fig. 11 shows the thermal neutron distribution down the length of a thermal column (Stephens and Williams 1960, unpublished). Access ports to the internal region of the thermal column are usually available. These can be used for placing specimens in the column (down the length of the column a wider choice of relative radiation intensities is available) or they can pass

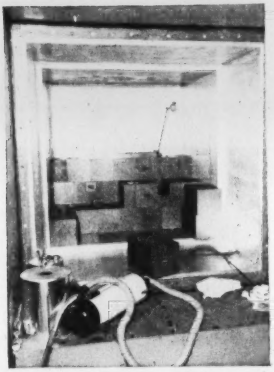


Fig. 10. A thermal column being filled with graphite. In order to minimize impurities which might absorb neutrons clean conditions were maintained during the loading procedure. With the graphite removed the cavity could be used as a large area source.

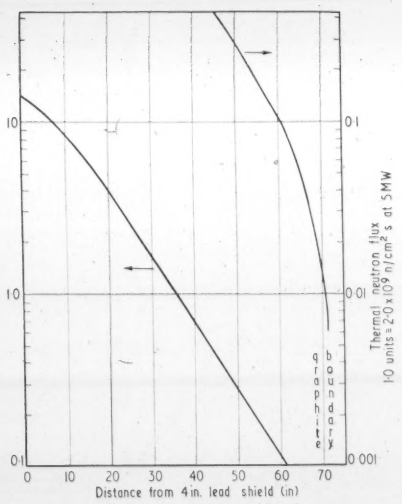


Fig. 11. The thermal neutron flux distribution along a thermal column (Stephens and Williams 1960, unpublished).

thermal neutron beams into the experimental area. Large objects can also be irradiated in the thermal column.

5.5. *Large area sources*

In some types of research, such as shielding studies and metallurgy, the experiments to be irradiated may be large, say three to six feet square. In order to accommodate such units in, or partly in, a high radiation field large area sources are being built into reactors. The facility may penetrate the shield right up to the core if a fission spectrum may be required. The radiation levels at the inside surface may be almost the same as that at the surface of the core itself. In the pool type of reactor the whole facility may be a tunnel through the water from the core region to a hot cell where the experiment can be unloaded from its trolley. Shielding experiments can be exceedingly heavy and the trolley must be able to transport them, one with its load may weigh up to 100 tons.

5.6. *Fast transfer units*

Fast transfer units are devices for the rapid insertion and removal of a specimen from the radiation field. They are usually pneumatically operated. The specimen is contained in a can, usually of aluminium or a plastic material to reduce induced activity, although super-pure iron capsules have been used. This can moves in a tube which passes through the reactor and terminates either at the reactor top or in one of the neighbouring laboratories. Such facilities are useful in work with short-lived radio-isotopes—to the physicist studying their behaviour, to the chemist using them in activation analysis, or to the physician requiring them in therapy.

Most research reactors, of 1 MW or more, have one or more fast transfer units. In the BR1 graphite moderated reactor at Mol, Belgium, the aluminium can is moved by high pressure carbon dioxide and the removal time can be less than 0·1s. The thermal and epithermal neutron fluxes at the irradiation point are both about 10^{12}/cm^{2}s.

Another means of producing a short irradiation is to pulse the reactor output; this is discussed below.

5.7. *Used fuel elements*

Often the fuel has to be stored for some time, after being used in the reactor, before it is returned for chemical processing. The γ-radiation from the fission products may be useful for certain experiments. The fuel is usually stored in water, a depth of 20 ft is typical. The intensity obtained will depend on the power density in the particular reactor. The γ-irradiation facility of the ORR (Oak Ridge Reactor) which operates at a reactor power of 20 MW produces fields up to 10^7 r/h. If the fuel is in water then there is a small neutron flux due to the D(γ, n) reaction, from the deuterium in normal water. This amounts to about 10^4 thermal neutrons/cm^{2}s for a γ-ray intensity of 10^7 r/h; it can be reduced to insignificant proportions by surrounding a region with a good neutron absorber, say cadmium.

6. Special applications

6.1. *Pulsed reactors*

Reactors can be designed to produce intense bursts of radiation. At one end of this scale is the small, fast reactor designed for a single pulse only—the atomic bomb. At the other end is a relatively simple research reactor which could belong to a large university. The use of pulsed reactors for general research is still in its infancy; the work will not always be simple, working in a restricted time scale, but the results should be worth while.

The research programmes with such reactors cover many topics, for example the delayed radiations from fission, short-lived isotopes, the relationship between dose rate and the biological effects of neutrons, γ-rays and charged particles,

transient radiation effects on materials and the examination of monitor response in criticality accidents. Further possible topics of research include the kinetic behaviour of reactors, the testing of reactor safety devices, activation analysis using short-lived activities (with the possibility of making analyses at various times on the same specimen), high temperature effects, the properties of neutrinos and other experiments in nuclear physics and biology. These research applications are discussed further by McReynolds (1960). The performance of several types of pulsed fast reactor is given in Table 5. Several pulsed fast reactors have been operating, mainly at Los Alamos, the bare (unreflected) reactor called GODIVA is typical of this type. These provide short pulses, of the order of 10–100 μs long, but it is unlikely that such reactors will be generally available for research in the near future.

Table 5. *Characteristics of pulsed neutron reactors*

	GODIVA	KEWB	SPERT	TRIGA	TREAT
Type, as in Tables 1 and 2	Fast	10	6	8	11
Energy (MJ)	0·8	8·5	15	24	950
Neutrons	$6{\cdot}8 \times 10^{16}$	$6{\cdot}5 \times 10^{17}$	$1{\cdot}2 \times 10^{18}$	$1{\cdot}9 \times 10^{18}$	$7{\cdot}5 \times 10^{19}$
Excess reactivity above prompt critical (%)	0·05	~6	1·6	1·5	4·2
Initial period (ms)	0·012	1	4	3	40
Peak power (GW)	10	2	1	2	7
Pulse width at half maximum (ms)	0·040	3·5	15	10·5	140
Average temperature rise (°C)	100	75	75	300	260
Interval between pulses	~1 h	~15 min	~15 min	~10 min	~4 h

If a pulse of neutrons is required with a width of the order of one microsecond or less then accelerators have to be used, or atomic bombs. Hughes proposed that small atomic bombs could be used for nuclear spectroscopy, which is concerned with neutrons in the energy range up to a few kilo-electron volts. Hughes (1960) estimated that one 10 kT explosion would produce a number of neutrons equivalent to those obtained by operating a high intensity pulsed reactor (10^{17} neutrons/s in a 10 μs pulse every 0·1 s) for 3000 years.

6.2. *Power reactors*

The testing of prototype fuel elements in large quantities has been carried out in the power reactors at Calder Hall and Chapelcross. The United Kingdom Atomic Energy Authority Annual Report for 1959 stated that these tests have indicated faults which would have led to large failure rates in operation. However, another main purpose of using power reactors as a radiation source is to ensure that adequate numbers are tested. The failure rate of the standard design of Calder Hall fuel element is less than 0·1%; indeed if the reactor is to operate successfully such low failure rates are necessary. Thus, it is essential that large numbers of a prototype fuel element are tested in order to detect effects which are significant at this level. A large power reactor is the only place where testing in adequate quantity can be carried out.

6.3. *Medical uses*

X-rays have been used for over half a century for therapy. Fast and thermal neutrons have been also used since about 1938. In those early days the neutrons were produced by cyclotrons and such sources are still in use. However, the very high neutron fluxes available from reactors are now being used; indeed six reactors, with powers in the range 5 W to 5 MW, are being used entirely for medical or associated biological research. They are attached to hospitals or medical research units in America. Several other research reactors have been designed with special facilities for medical research.

Irradiation with thermal neutrons can be used after administration of a stable isotope possessing a high thermal neutron capture cross section. By proper choice of the isotope the resultant reaction will produce ionizing particles of very short range, thus localizing the heavy damage. Also, if the isotope is chosen correctly it will be preferentially situated in the region requiring treatment. An example of this technique of neutron capture therapy is the treatment of brain tumours, using the $^{10}B(n,\alpha)$ reaction to produce the desired radiation (Brownell and Street 1958). Both the 30 MW general purpose Brookhaven reactor and the M.I.T. reactor have been used in the work. Thermal neutron fluxes of the order of 10^9 to 10^{10}/cm^{2}s are typical. With the latter flux 15 μg of ^{10}B per gramme of tissue would produce a dose rate of 100 r/min. Neutron doses of about 10^{12}/cm^2 have been used (Farr *et al.* 1958). Fast neutron therapy was used about 1938, but after-effects were unsatisfactory; however, there may be further research into this method of treatment.

The M.I.T. reactor operates at a maximum power of 1 MW. Its medical facility is a large vertical port from the reactor tank to a room beneath. Three layers of bismuth below the core attenuate the γ-radiation. These are followed by a beam shutter consisting of a tank of water; this is emptied before the facility is used otherwise the water absorbs most of the neutrons. Below this tank is a boron shutter which, when closed, will attenuate any remaining neutrons. Finally, there is a thick lead and bismuth shutter further to reduce γ-radiation. The design is such that a thermal neutron flux of 4×10^{10}/cm^{2}s should be available, with a background of 10^7 epithermal neutrons/cm^{2}s and about 100 r/h of γ-radiation. The room below is a complete hospital operating room. The doctor in charge of the treatment has control of the facility and can shut down the reactor if necessary. Fast transfer units are also being installed in the reactor for the production, in amounts up to a few curies, of short-lived isotopes for medical use. The units are also used to simplify the production of other isotopes, such as ^{165}Dy, required for internal therapy (G. L. Brownell 1960, unpublished).

The Brookhaven Medical Reactor has a maximum operating power of 5 MW. There are two treatment rooms each separated from the reactor by a 16 in. square port closed by a vertical shutter. The radiation intensities available at the patient's head are 10^{11}/cm^{2}s (thermal), about 10^4/cm^{2}s (fast) and a γ-radiation flux of about 10^4 r/h. There is also a large area source at the thermal column, the maximum thermal neutron flux is about 10^{10}/cm^{2}s and the intensities of the various radiations can be varied by means of filters.

Acknowledgments

The author is indebted to Dr. T. E. Allibone for permission to publish this paper, to Mr. E. Hesse and Miss M. Cresswell for assistance with Table 3, and to Mr. K. W. Hetzel for information on the TRIGA type of reactor. Fig. 7 is shown by courtesy of the United Kingdom Atomic Energy Authority.

References

BROWNELL, G. L., and STREET, W. H., 1958, *Proc. of U.N. Conference on the Peaceful Uses of Atomic Energy*, **26**, 444.

CHASTAIN, J. W., 1958, *U.S. Research Reactor Operation and Use* (Cambridge, Mass.: Addison-Wesley).

COURT, A. J., and DOWNES, K., 1960, *Trans Amer. Nucl. Soc.*, **3**, 1, 10–5.

FARR, L. E., *et al.*, 1958, *Proc. of U.N. Conference on the Peaceful Uses of Atomic Energy*, **26**, 451.

GLASSTONE, S., 1956, *Principles of Nuclear Reactor Engineering* (London and New York: Macmillan).

GLASSTONE, S., and EDLUND, M. C., 1952, *The Elements of Nuclear Reactor Theory* (New York: Van Nostrand).

HUGHES, D. J., 1953, *Pile Neutron Research* (Reading, Mass.: Addison-Wesley).

—— 1960, *Nucleonics*, **18**, No. 7, 77.

MCREYNOLDS, A. W., 1960, *General Atomics Report*, GA 1055.

WESTCOTT, C. H., WALKER, W. H., and ALEXANDER, T. A., 1958, *Proc. of U.N. Conference on the Peaceful Uses of Atomic Energy*, **16**, 70.

Notes and comments

New scientific monographs

Elek Books have started a new series of scientific monographs bearing the imprint Logos Press. The consultant editor is Professor Sir Harrie Massey and the executive editor is Dr. A. P. Willmore.

The Wave-guide Mode Theory of Wave Propagation by Dr. K. G. Budden will be the first monograph to appear in this new series. The book deals with the mathematical theory of two problems: the propagation of radio waves to great distances over the surface of the earth and the propagation of sound waves to great distance under water. The book will be of interest to graduate students of physics, mathematics and communication engineering, particularly those concerned with the use of low frequency radio waves for communication and navigational aids, to meteorologists concerned with electrical noise, and to acoustical engineers and others interested in underwater sound.

This book will be published in October by Logos Press, Scientific Publications (Elek Books Limited), 14 Great James Street, London, W.C.4.

The properties and applications of 'K' Monel alloy

We have received a copy of a new publication on the properties of 'K' Monel nickel-copper alloy. It consists mainly of charts and graphs on the physical and mechanical properties of the alloy with sections on the available forms and applications.

The publication is obtainable free on request from the Publicity Department, Henry Wiggin & Company Limited, Thames House, Millbank, London, S.W.1.

Chemical and Petroleum Engineering Exhibition

The Second Chemical and Petroleum Engineering Exhibition will take place at Olympia, London, from 20th–30th June 1962, under the joint sponsorship of the British Chemical Plant Manufacturers Association and the Council of British Manufacturers of Petroleum Equipment.

The Third Congress of the European Federation of Chemical Engineering will be held during the course of the exhibition with symposia on Process optimization, Interaction between fluids and particles, The handling of solids and The physics and chemistry of high pressures.

The Fourth Congress of the Federation of European Petroleum Equipment Manufacturers (Fédération Européene des Constructeurs d'Equipement Petrolier) will also be held during the exhibition from 25th–28th June at Church House, Westminster. Approximately 20 technical papers, covering Automation, New corrosion resistant materials, Temperature insulation, Aircraft refuelling and Future trends in the design of drilling equipment for the petroleum industry, will be presented.

Further information may be obtained from The British Chemical Plant Manufacturers Association, 14 Suffolk Street, London, S.W.1, or from The Council of British Manufacturers of Petroleum Equipment, 2 Princes Row, London, S.W.1.

Principles governing the viscous flow of suspensoid sols

by T. V. STARKEY, B.Sc., Ph.D., F.Inst.P. (§1), V. A. HEWLETT, M.Sc., F.R.I.C. (§2), J. H. A. ROBERTS, B.Sc., A.R.I.C. (§3), and R. E. JAMES, M.Sc., A.Inst.P. (§4). The Technical College of Monmouthshire, Crumlin, Mon.

MS. received 16*th February* 1961, *in revised form* 15*th May* 1961

Abstract

A theory relating to the viscous flow of stable suspensoid sols, and briefly outlined in earlier papers is extended in the present work.

It suggests (i) *that concentration patterns are developed within suspensoid sols as a result of shearing*, (ii) *that these patterns arise through the operation of 'least action forces'. (The existence of such forces is required in order that the principle of least action shall be applicable to those changes in the configuration of the system of particles comprising the disperse phase of the sol which accompany shearing)*, (iii) *that the velocity distribution pattern within such a sol is related in any given case to the concentration pattern*, (iv) *that the anomalous properties of suspensoid sols under shear arise in consequence of the development of these patterns, and* (v) *that changes in these properties are due to changes in the parameters defining these patterns, and these, in turn lead to variations in the magnitudes of the least action forces with changes in the shearing conditions. It leads, in the case of flow through capillary tubes to the following conclusions:*

(1) *Shear dependence of viscosity is a universal characteristic of such sols.* (2) *The overall mean sol concentration within a tube during flow is always less than the concentration in the reservoir from which it is supplied.* (3) *The mean concentration over any cross section diminishes systematically along a tube from a maximum value equal to the reservoir concentration (near to the entrance) to a minimum value at points nearer to the exit.* (4) *The overall mean sol concentration varies with flow rate from a maximum value (equal to the reservoir concentration) at zero rate of flow to a minimum value (always greater than one half of the reservoir concentration) at the highest rate for non-turbulent flow.* (5) *The velocity profile, as defined by the ratio of mean to maximum streaming velocity, is also shear dependent, the ratio varying systematically between limits of one half and one.* (6) *The ratios of mean to maximum velocity and of mean tube concentration to reservoir concentration are equal to one another for sols showing no appreciable Brownian movement and flowing in very long tubes.* (7) *These velocity and concentration ratios can under the conditions mentioned in* (6), *be expressed as a simple function of the relative viscosity of the sol.* (8) *The axial sol concentration increases systematically along the tube from a minimum value (equal to the reservoir concentration), near the tube entrance, to a maximum value which may be many times as great at points nearer to the exit. (This hyperconcentration will under appropriate conditions lead to the formation of a thrombus.)*

Experiments designed for the investigation of the conclusions numbered (1), (2), (4), (5) *and* (7) *are also described. The results yielded by these experiments support the conclusions mentioned.*

In the light of these results Einstein's law which is based on the assumption of random particle distribution throughout a sol undergoing shear is seen to be fundamentally unsound.

1. Theoretical

1.1. *General introduction*

ALTHOUGH shear dependence of viscosity has been reported both for suspensions and macromolecular solutions on many occasions considerable doubt still appears to exist as to whether this phenomenon is, or is not, a general characteristic of suspensoid sols. If it is, then it is certainly fundamental to a complete understanding of the nature of viscous flow in such sols that it should be recognized. As an illustration of the present uncertainty on this question the conclusions of Richardson and Tyler (1933) and of Higginbotham, Oliver and Ward (1958) may be quoted.

Richardson and Tyler, using suspensions of rice starch grains in carbon tetrachloride and paraffin, obtained results from which they concluded that the coefficient of viscosity in a disperse system diminishes with increasing velocity gradient. Higginbotham, Oliver and Ward, on the other hand, found that the suspensions of spherical polymer particles in aqueous glycerol solutions which they used, behaved as Newtonian fluids. Since no fundamental difference appears to have existed between the sols employed it seems not unlikely that the discrepancy arises as a result of the manifestly different shearing conditions used in the two cases.

Similar uncertainty exists also in regard to the viscosity shear dependence of high polymer solutions. In connection with these Harland (1955) concludes that this phenomenon is appreciable only where the degree of polymerization exceeds a value of about 2000. Whilst this appears to be roughly true for measurements made under near-standard shearing conditions it may well prove to be far from true under appropriate non-standard conditions.

This general uncertainty regarding a very fundamental characteristic of suspensoid sols is matched by no less uncertainty in regard to its explanation. Thus, different authors have employed assumptions relating to the different, but particular, characteristics of the disperse phases of the sols with which they have been concerned, as a basis for explaining their viscosity shear dependence. These have included such assumptions as that the particles are deformed under shear, or that the particles are porous and that drainage results from shearing, or (in those cases where the particles are anisodiametric), that preferred particle orientation occurs and this, as with other characteristics mentioned, varies with shear rate (Immergut and Eirich 1953). Whilst such assump-

tions may seem reasonable enough with regard to the particular sols which the respective authors have used, they can have no general validity. And further, none of the assumptions referred to can account for the variations of viscosity with shear rate reported for suspensions of non-porous and roughly spherical rigid particles (Richardson and Tyler 1933).

In earlier papers (Starkey 1955, 1956) a flow mechanism was described in accordance with which these viscosity variations were explained in terms of changes occurring, with shear rate, in the positions of the particles in the velocity field. It was pointed out that this mechanism would provide an explanation of non-Newtonian behaviour independent of any assumed attributes particular to individual suspensoid sols, and therefore, of general applicability. This mechanism depends on the fact that the particles comprising the disperse phase of the sol will follow paths during shear such that the resulting change in the configuration of the system is effected in accordance with the requirements of the principle of least action. This can readily be seen to imply the existence of forces acting on the particles of the disperse phase displacing them from the streamlines on which they initially move to others where the velocity gradient is lower. The combined net effects of these forces will be to develop a *concentration pattern* within the sol such that maximum concentration occurs where the velocity gradient is least. (In what follows these forces will be referred to as 'least action forces'.)

As the values of the least action forces, and consequently of the parameters defining the pattern, would be expected to vary with shear rate, and since it is inconceivable that other forces always equal and opposite to these can exist in suspensoid sols under shear, this mechanism requires the shear dependence of viscosity as a characteristic of all suspensoid sols. The extent to which this occurs, however, would be expected to vary from case to case in a manner depending not only on the initial particle concentration, size and shape, but also on the geometry of the shearing system and the range of shear rates used.

1.2. *The laminar flow of suspensions in tubes. (Development of the theory)*

The theory dealt with here was briefly outlined in earlier papers (Starkey 1955, 1956). In accordance with the mechanism indicated there, suspended particles, although distributed randomly over the whole section of a tube on entry to it, would be expected as a result of the operation of the least action forces to move towards the axis as they progress along the tube, so forming a 'core' of suspension around the tube axis, surrounded by a 'sheath' of suspension fluid, devoid of particles, adjacent to the wall. (The fully developed core, it should be noted, is not produced immediately at the tube entrance, and in short tubes the phenomena dealt with in this paper may be partially obscured by this 'entrance effect'.)

Consider the stability of this core when Brownian movement is (i) negligible and (ii) finite.

(i) In this case no forces which could cancel the effects of the least action forces mentioned can be conceived, and in a long tube the latter would be expected to operate continuously until all particles had been displaced laterally into a region of zero velocity gradient where particle rotations would cease and the least action forces would consequently vanish.

(ii) Where Brownian movement is not negligible diffusion would arrest the process described in (i) before it is complete.

A suspensoid sol which shows no Brownian movement, flowing under steady conditions through a long tube (that is, one in which the entrance effect mentioned earlier can be neglected), can therefore be regarded as consisting of a 'core' of sol of radius ρ, surrounded by a 'sheath' of suspension fluid of thickness $R - \rho$, where R is the tube radius, such that velocity gradients in the core are negligible compared with those in the sheath. For extremely dilute sols the core would be expected to assume the form of a 'vortex street' as mentioned in an earlier paper (Starkey 1955).

Let us now consider quantitatively some consequences of this particle distribution for a suspension of particles, dispersed in a perfectly homogeneous liquid, maintained at constant temperature and flowing slowly through a long capillary tube.

1.2.1. *Velocity distribution within the 'sheath'.*

Adopting the method of analysis commonly employed in the derivation of the Poiseuille formula (Worsnop and Flint 1931) the net force acting on a cylindrical element of radius r (where $r > \rho$), thickness δr and length l is seen to be equal to

$$-(dF/dr)\delta r = 2\pi r \delta r P \tag{1}$$

where F is the shearing force acting on one of the cylindrical surfaces of the element and P is the pressure difference under which flow takes place. Since

$$F = \eta 2\pi r l dv/dr \tag{2}$$

where η is the viscosity of the suspension fluid and v the velocity at $r = r$,

$$-\frac{d}{dr}\left(2\pi r l \eta \frac{dv}{dr}\right) = 2\pi r P \tag{3}$$

from Eqn (1).

By integrating (3) twice, and since $dv/dr = -P\rho/2\eta l$ when $r = \rho$, and $v = 0$ when $r = R$, it can be shown that

$$v = \frac{P(R^2 - r^2)}{4\eta l} \tag{4}$$

1.2.2. *Volume emerging per second.*

Using the method referred to in §1.2.1, the volume V emerging per second from the tube is seen to be

$$V = \int_{\rho}^{R} 2\pi r v dr + \frac{\pi \rho^2 P(R^2 - \rho^2)}{4\eta l}. \tag{5}$$

By substituting for v from Eqn (4), integrating and simplifying

$$V = \frac{P\pi(R^4 - \rho^4)}{8\eta l}. \tag{6}$$

If ρ becomes zero, Eqn (6) reduces to the Poiseuille formula,

If V varies non-linearly with P, and therefore the viscosity of the sol is shear dependent, $\pi(R^4 - \rho^4)/8\eta l$ must vary with P; but since ρ is the only quantity in this expression which can vary with P under the given conditions, viscosity shear dependence must be due to changes in core radius with rate of flow.

1.2.3. *Relation of relative viscosity to core and tube radii.*

If η_a is the apparent viscosity of the sol, then by definition

$$\eta_a = \frac{P\pi R^4}{8Vl},$$

but, from Eqn (6)

$$\eta = \frac{P\pi(R^4 - \rho^4)}{8Vl},$$

therefore the relative viscosity η_{rel}, is given by

$$\eta_{rel} = \frac{\eta_a}{\eta} = \frac{R^4}{(R^4 - \rho^4)} \tag{7}$$

or

$$\rho^4 = \frac{R^4(\eta_{rel} - 1)}{\eta_{rel}}. \tag{8}$$

As seen from Eqn (8) a diminution in relative viscosity with (say) increased rate of flow is associated with a diminished core radius.

1.2.4. *Relation of concentrations and velocities in the sol.*

Let a suspensoid sol of particle concentration c_0, enter a long capillary tube at a mean velocity of v_m. The mass of particles entering per second is, therefore $\pi R^2 c_0 v_m$.

Let the mean concentration in the tube be c_t. Since all particles will on emergence be in the core, and will move with the core velocity v_ρ, the mass of particles leaving per second will be $\pi R^2 c_t v_\rho$.

Since the masses entering and leaving per second, under steady state conditions, are the same

$$c_0 v_m = c_t v_\rho. \tag{9}$$

Since $v_\rho > v_m$, $c_t < c_0$; the mean concentration in the tube is, therefore, always less than the concentration in the reservoir.

1.2.5. *Relation of velocity, viscosity and concentration ratios.*

From Eqn (4), the maximum velocity v_ρ is given by

$$v_\rho = \frac{P(R^2 - \rho^2)}{4\eta l} \tag{10}$$

and from Eqn (6) the mean velocity v_m by

$$v_m = \frac{V}{\pi R^2} = \frac{P(R^4 - \rho^4)}{8\eta l R^2} \tag{11}$$

whence

$$\frac{v_\rho}{v_m} = \frac{2R^2}{R^2 + \rho^2}. \tag{12}$$

From Eqns (8) and (12), therefore

$$\frac{v_\rho}{v_m} = \frac{2\sqrt{\eta_{rel}}}{\sqrt{\eta_{rel}} + \sqrt{(\eta_{rel} - 1)}} \tag{13}$$

$$= \frac{c_0}{c_t} \text{ from Eqn (9).}$$

If suspensoid sols are then viscosity shear dependent they would be expected also to be both velocity-ratio shear dependent and concentration-ratio shear dependent. Further since as $\eta_{rel} \to \infty$, $v_\rho/v_m \to 1$ and as $\eta_{rel} \to 1$, $v_\rho/v_m \to 2$ these velocity and concentration ratios would always be expected to lie between the limits 1 and 2.

Equation (13) also suggests that if relative viscosity is made to vary by any means whatsoever, e.g. by changing rate of flow, initial sol concentration, geometry of apparatus, etc., the velocity and concentration ratios will vary accordingly.

1.2.6. *Core concentrations.*

The mass of particles leaving the tube per second under steady state conditions is $\pi\rho^2 c_\rho v_\rho$ where c_ρ is the core concentration. If this is equated to the mass $\pi R^2 c_0 v_m$ entering per second,

$$\frac{c_\rho}{c_0} = \frac{R^2 v_m}{\rho^2 v_\rho}. \tag{14}$$

From Eqn (12), therefore,

$$\frac{c_\rho}{c_0} = \frac{R^2 + \rho^2}{2\rho^2}. \tag{15}$$

From Eqn (8) this becomes

$$\frac{c_\rho}{c_0} = \frac{\sqrt{\eta_{rel}} + \sqrt{(\eta_{rel} - 1)}}{2\sqrt{(\eta_{rel} - 1)}} \tag{16}$$

In accordance with Eqn (16), and since η_{rel} is always greater than unity, the core concentration is always greater than the reservoir concentration. Further, any diminution in the relative viscosity is, in accordance with this same equation, accompanied by an increase in core concentration relative to reservoir concentration. Under conditions of high initial concentration and high flow rate therefore, the core concentration will become very great indeed, and under extreme conditions so much so that particle contacts are inevitable.

This hyperconcentration of the sol in the axial region of a tube can thus give rise under appropriate conditions to the development of a thrombus as in coronary thrombosis.

1.2.7. *Relation between maximum and mean concentrations (hyperconcentration).*

From Eqns (12) and (13)

$$\frac{c_t}{c_0} = \frac{R^2 + \rho^2}{2R^2}. \tag{17}$$

From Eqns (15) and (17) it follows that

$$\frac{c_\rho}{c_0} = \frac{c_t}{2c_t - c_0}. \tag{18}$$

It is seen from Eqn (18) that as $c_t/c_0 \to 1$, $c_\rho/c_0 \to 1$ also, and as $c_t/c_0 \to \frac{1}{2}$, $c_\rho/c_0 \to \infty$. That is, for a given initial concentration the axial concentration increases with decreased relative viscosity, whilst the mean concentration within the tube simultaneously diminishes.

The theoretical relationships derived are illustrated in Fig. 1.

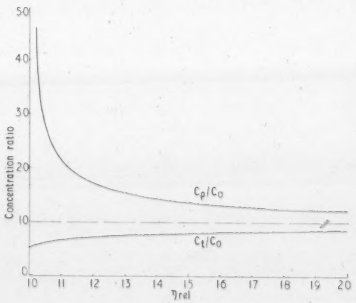


Fig. 1. Theoretical curves showing variation in ratios of the mean and maximum 'tube' concentrations to the initial concentration, with relative viscosity.

1.2.8. *Entrance effects.*

Near to the entrance of all tubes the concentration pattern described at the beginning of §1.2 is only in the process of being developed. In this region variations in concentration from point to point along the tube must be expected.

Thus, if the mean streaming velocity of the particles in any disk-shaped element of the tube is v_p, it can be shown by the repeated use of the argument employed in the derivation of Eqn (9), (noting that this applies with equal validity either to the element or the whole tube) that the mean concentration, c_e, of the sol within that element, is related to the initial concentration by

$$c_e = \frac{c_0 v_m}{v_p}. \tag{19}$$

Since v_p increases systematically along the tube from a minimum value, v_m, at the entrance, to a maximum value v_p where the fully developed core starts, c_e diminishes correspondingly from c_0 to $c_0 v_m/v_p$. Thereafter it remains constant at the latter value.

If the maximum concentration in the same disk-shaped element is c_p, it can be shown, similarly, that

$$\frac{c_p}{c_0} = \frac{c_e}{2c_e - c_0}. \tag{20}$$

From Eqn (19) therefore

$$\frac{c_p}{c_0} = \frac{v_m}{2v_m - v_p}, \tag{21}$$

but since, as just mentioned, v_p increases along the tube, c_p must also increase. The minimum value c_0 of c_p will, as seen from Eqn (21) occur at the entrance, where $v_p = v_m$. As v_p increases, however, c_p will increase correspondingly, reaching its maximum value $c_0 v_m/(2v_m - v_p)$ at the point at which the mean concentration c_e has diminished to its minimum value; thereafter it will remain constant.

For this reason a thrombus would not be expected to form at the entrance to a coronary artery, but rather to develop after the point at which c_p has reached that value at which corpuscles come into contact with one another.

In the light of these concentration variations, measured values of the overall mean tube concentration would be expected to be higher when the entrance effects are not negligible, for the process of measuring them involves averaging over the whole tube length, mean concentrations which, up to a point, diminish along the tube.

Measured values of viscosity would also be expected to be higher under these conditions, for the sol in the entrance region is only in the process of assuming its least action state.

The ratio of the mean to the maximum streaming velocity, on the other hand, would be expected to be lower when entrance effects are significant, for the concentration pattern would be less 'plug-like' under these conditions.

In any given case, and since an increase in the least action forces gives rise to reduced apparent viscosity, the entrance effect can be diminished by any device leading to diminished apparent viscosity. A reduction in tube bore (Higginbotham, Oliver and Ward 1958), other parameters being unaltered, would consequently be expected to reduce the entrance effects.

Again, and remembering that the least action forces vary with particle size and shape characteristics (Starkey 1955, 1956), the core formed by shearing a sol containing particles of various sizes and shapes (and similarly the thrombus formed in cases of high hyperconcentration), would be expected to exhibit a characteristic arrangement of the particles composing it. Thus, for a sol containing particles all of the same shape but of several different sizes, a section through the core (or through the thrombus), normal to the direction of flow, would be expected to display a concentric ring pattern in which the largest particles would be concentrated in the inner rings.

A review of related researches has recently been published by G. W. Scott Blair (1958). A list of references to the relevant literature is given at the end of this review.

2. Shear dependence of viscosity of suspensoid sols

In terms of the mechanism under investigation, and outlined in the previous section, variations of viscosity with flow rate are fundamentally the result of variations in the magnitudes of least action forces, for an increase in the rate of flow of a sol through a capillary tube would be expected to give rise to increased least action forces, these to diminished core radius and this to reduced apparent viscosity.

As mentioned previously, however, the formation of a fully developed core assumes both the absence of appreciable Brownian movement and shearing in a long tube. The effects described would consequently be expected to be very small indeed for sols whose disperse phases are molecular and especially so for shearing in relatively short tubes.

Nevertheless, since solute molecules constitute discontinuities in a solution, shearing would be expected to give rise to the formation of vortices (Starkey 1955) and so to the least action effects described. Even true solutions then would be expected to show some viscosity shear dependence, however small, under appropriate conditions, and this, moreover, would be expected to involve viscosity variations with rate of flow of the same kind as is met in markedly shear dependent sols, that is, the apparent viscosity would be expected to decrease with increased flow rate.

Viscosity-rate of shear curves for those suspensoid sols which have been shown to be viscosity shear dependent, indicate a rate of fall of viscosity which decreases as the rate of flow is increased. Thus, even in the case of sols showing marked viscosity shear dependence only small viscosity changes are to be expected in the higher shear-rate range. Since the present experiments are intended to investigate viscosity shear dependence in sols which are commonly regarded as Newtonian it is immediately apparent that for optimum conditions the lowest shear-rate range attainable should be used. An exponential flow type of apparatus was consequently employed in the present investigations since it is possible by this means to deal with fluids under extremely low rate of flow conditions. The apparatus and its method of use are described below.

2.1. *Apparatus and method*

The apparatus used is illustrated to scale in Fig. 2. It comprises a U tube viscometer, constructed of precision bore tubing and having a capillary of length 12 cm and internal diameter $0{\cdot}050 \pm 0{\cdot}001$ cm, in one limb, a 1 in. micrometer screw gauge reading to 0·002 in., and a thermostatically controlled water bath. Both the viscometer and the micrometer are attached through levelling and clamping screws (not shown) to a bridge fixed to the water bath. By this means the micrometer and the capillary tube can be made coaxial and vertical and rigidly fixed in relation to one another. A platinum wire probe, sealed into a glass tube which passes through a guide ring in the bridge, is fixed to the micrometer head; it can consequently be moved vertically relative to the free surface of liquid contained in the viscometer by rotating the head.

The experimental procedure involves the following stages. After filling the viscometer with the liquid to be tested, it is

immersed in the water contained in the thermostatic bath, the temperature of which can be controlled to 0·01 deg C. After levelling the viscometer the liquid in it is allowed to approach its equilibrium position until the remaining head

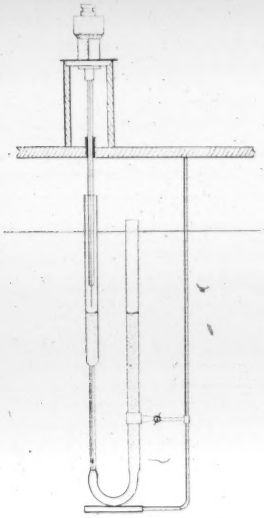


Fig. 2. Low head viscometer.

becomes too small to be measured by the micrometer, i.e. it is less than 0·002 in. Repeated readings of the micrometer corresponding with the break of the probe away from the free surface of the liquid are then taken; the zero reading is obtained from these.

The liquid under test is now displaced by means of a hand bulb so as to establish a fixed initial pressure head of 0·8250 in. It is then allowed to flow under the progressively diminishing pressure and the times of flow, from the instant at which the head is 0·8250 in., corresponding to the break of liquid away from the probe, are recorded for each successive complete revolution of the micrometer. These time intervals are measured by two stop watches previously calibrated against a National Physical Laboratory standard.

Under the conditions described the instantaneous head h is, for a Newtonian liquid, exponentially related with the cumulative time t. That is

$$h = h_0 e^{-t/\alpha} \qquad (1)$$

where h_0 is the value of h at $t = 0$, i.e. 0·8250 in., and α is a constant which, under the given conditions, is proportional to the viscosity of the liquid.

For a non-Newtonian liquid α varies with h; it increases as the apparent viscosity of the liquid increases.

In the present series of experiments, values of α, calculated from the logarithmic form of Eqn (1) and the corresponding values of h, were obtained at a temperature of 25·00° C for each of a number of liquids selected so as to include solutions of relatively low molecular weight, polymers on the one hand and of non-associated liquids on the other.

2.2. *Experimental results*

Typical results are shown in Fig. 3, the individual curves of which relate to A, carbon tetrachloride, B, 90% solution of phenol in water, C, 2% solution by weight of nylon

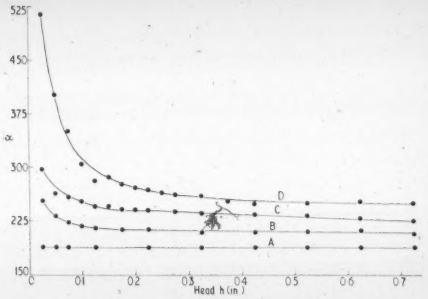


Fig. 3. Variation of α (and hence of apparent viscosity) with applied pressure for A, carbon tetrachloride ($\alpha = 50$); B, 90% phenol/water solution ($\alpha = 1850$); C, 2% solution of nylon (average m.w. 3876) in 90% phenol/water ($\alpha = 3650$); D, 2% solution of nylon (average m.w. 14 900) in 90% phenol/water ($\alpha = 7490$).

(average m.w. 3876) in 90% phenol/water and D, 2% solution by weight of nylon (average m.w. 14 900) in 90% phenol/water. Of these carbon tetrachloride alone behaves as a Newtonian fluid throughout the entire range of pressures used.

Whilst the other liquids are approximately Newtonian at the high pressure end of the range, all show progressive increases in apparent viscosity as the pressure is reduced. These increases follow the pattern commonly reported for markedly shear dependent suspensoid sols, i.e. the rate of increase of viscosity rises as the applied pressure diminishes. These facts are all consistent with the mechanism outlined in §1.

3. Concentration phenomena in suspensoid sols under shear

The flow mechanism under consideration requires that the mean concentration of a suspensoid sol within a capillary tube through which it is flowing, shall always be less than that in the reservoir from which it is derived. Further, this mean concentration must, for a given sol, diminish relative to the concentration in the reservoir as the rate of flow is increased.

In accordance with this mechanism the mean velocity of the sol particles at the exit to the tube must be greater than that at the tube entrance since particles are in general laterally displaced from lower to higher velocity streamlines whilst within the tube. Since the mass of sol flowing per second and the mean sol concentration under steady state conditions are the same at exit and entrance, and yet the mean particle velocity is greater at exit than at entrance, the effluent sol must have been derived from a less concentrated source than the entering sol. That is, the mean concentration within the tube must be less than the reservoir concentration. And further, since the mean particle velocity at the exit must be greater than at the entrance to an extent which increases with increases in the values of the least action forces at higher flow rates, the mean tube concentration must diminish relative to the reservoir concentration as the rate of flow is increased.

This phenomen of concentration shear dependence was

dealt with mathematically in relation to the viscosity shear dependence associated with it, in §1 of this paper (see Eqn (13)).

The experiments described in the present section were designed to investigate the validity of this predicted concentration shear dependence.

3.1. *Apparatus and method*

The apparatus used in the present experiments is illustrated in Fig. 4. It consists essentially of a ten-litre storage reservoir

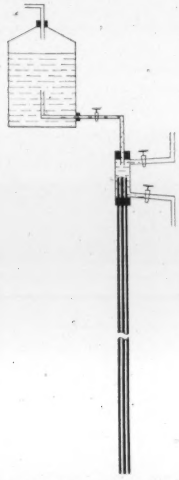


Fig. 4. Apparatus for the measurement of mean 'tube' concentrations during flow.

fitted with a magnetic stirrer and connected via a wide bore tap to a small subsidiary reservoir from which three almost identical capillary tubes project vertically. The subsidiary reservoir was provided with inlet and drainage tubes which could be closed by taps. The capillaries were 153 cm long and had a combined internal volume of 17·06 cm^3. The upper ends, both of the tube connecting the reservoirs, and the capillaries, were fixed near to the centres of the reservoirs.

A mercury manometer and either an exhaust or a compression pump (not shown in the diagram) were connected to the storage reservoir. By this means measured positive or negative pressures could be applied to the free surfaces of the suspension contained in this reservoir.

The suspension used in these experiments consisted of polystyrene beads in a glycol–water solution of the same density, to which 1 cm^3 of Teepol was added for each 100 g of beads used. The initial concentration of the suspension, obtained by a method to be described later, was $1{\cdot}52 \pm 0{\cdot}08$ g of beads per 100 cm^3 of suspension. The size range of the beads, obtained from standard sieves, was 251 μ to 422 μ.

For experiments of the present kind a very high degree of suspension stability is necessary and the following precautions were taken to ensure as far as possible that this requirement was met. The suspension was made by shaking together the sieved beads and the water–glycol solution, which had previously been made up to approximately the same density as the beads. Afterwards a few drops of glycol or water were added until no visible sedimentation occurred. If a sample of this suspension placed in a measuring cylinder adjacent to the storage reservoir showed no flocculation or sedimentation after standing for about an hour, the suspension was regarded as sufficiently stable for the present purpose. Arrangements were made so that as soon as this condition was achieved the main part of the experiments could be carried out within a relatively very short time. Under these conditions, and in view of the fact that a large volume of suspension was employed, the temperature within the latter remained very closely constant at 19·0° C throughout the experiment.

The experiments involved the measurement of the concentration of the suspension in the reservoir and the mean concentration of the suspension flowing in the capillaries under each of a series of applied pressure differences. The former concentration was measured as follows. After adjusting the pressure to some given value the suspension was allowed to flow to waste through the capillaries. When steady state conditions had been obtained a 100 cm^3 sample of the effluent suspension was collected in a standard flask such that when completely full it contained the volume specified. The suspension so collected was then filtered through a weighed Gooch crucible and the beads, after washing several times with water, were dried to constant weight in an air oven maintained at 60° C to 70° C. The reproducibility of the weight of the dry beads obtained provided an indication both of the stability of the suspension and of the existence of the steady state conditions referred to. The concentration of the suspension in the reservoir was calculated from these data.

Immediately after collecting the sample of effluent suspension, flow was suddenly stopped by closing the lower ends of the capillaries. The supply from the storage reservoir was also cut off and the subsidiary reservoir was emptied through the drainage tube. The contents of the completely filled capillaries were then allowed to run into a flask and the tubes were washed several times with water into the same flask. So as to increase the accuracy of the determination this process was carried out six times under the same conditions, and, after washing and drying the beads collected as before, the weight of the beads contained in the suspensions of the eighteen tubes was found. From this data, knowing the total internal volume of the capillaries, the mean concentration of the suspension flowing in them, and hence the ratio of this mean concentration to the initial suspension concentration, was determined.

The whole experiment was then repeated at each of a series of applied pressures.

3.2. *Experimental results*

Values of the concentration ratio varied in the manner illustrated in Fig. 5 between pressures of -4 and $+24$ cm Hg. Within this range the values of the ratio, as shown by the scatter of points about the curve, were reproducible to within about $\pm 1\%$. Outside this range, however, the variability in ratio values was appreciably higher than this. At pressures below -6 cm Hg there was evidence of flocculation within the tubes (though not in the reservoir) and correspondingly low ratio values, well below unity, were obtained, whilst at a pressure of about $+22$ cm Hg, turbulence set in.

The results obtained confirm the theoretical conclusions mentioned in the first paragraph of §3.

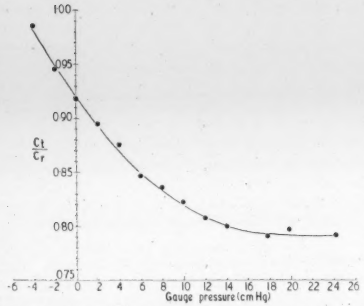


Fig. 5. Variation in mean 'tube' concentration with applied pressure.

4. Velocity profile phenomena in suspensoid sols under shear

As explained in §1 of this paper the operation of least action forces would be expected to lead to the development of a concentration distribution within a suspensoid sol flowing in a capillary tube, such that maximum concentration occurs in the axial region of the tube. The characteristic 'flattened paraboliform' velocity profile (Vejlens 1938) observable in such flow can readily be associated with this concentration distribution, for the central flattened region of the profile defines the 'core' in those cases where the 'entrance effect' is negligible.

Since changes in relative viscosity arise as a result of changes in core radius the velocity profile would be expected to vary with the relative viscosity. Thus, at very high relative viscosities, the associated core would be expected to extend nearly to the tube wall, and with it, the flattened region of the velocity profile. As viscosity is gradually reduced, however, the diameter of the core would progressively diminish, and the velocity profile would therefore approximate more and more closely to the parabolic form characteristic of the approximately Newtonian suspension fluid.

These velocity profile changes were considered in §1.2 where they were related to the associated changes in apparent viscosity and concentration. It was shown there that the ratio of the mean to the maximum velocity would be expected to increase between the limits of one half and unity as the relative viscosity is increased from unity to some very high value. The experiments described in the present section were designed to investigate these variations in the velocity ratios.

4.1. *Apparatus and method*

The apparatus used in the present experiments was designed for the simultaneous measurement of the viscosity and velocity ratios for each of a series of sols identical in all respects except concentration. It is illustrated in Fig. 6.

The apparatus consists essentially of a U-tube viscometer with one limb of length 117·5 cm and bore 0·139 cm, and the other of the same length but having a bore of 2·5 cm. These tubes are inserted into two identical reservoirs and are arranged so that their ends are near the centres of the reservoirs.

Identical floats, which rest on the liquid in these reservoirs, carry fine horizontal cross-wires. A plane mirror is fixed midway between these cross-wires so making it possible to observe the coincidence of one wire with the mirror image of the other without parallax. A jet, supplied from a separate adjustable reservoir, enters the U-tube near the lower end of the capillary limb. It is aligned with the axis of the capillary.

The sols used throughout consisted of polystyrene beads of diameters 422 μ and below, suspended in glycerol–water

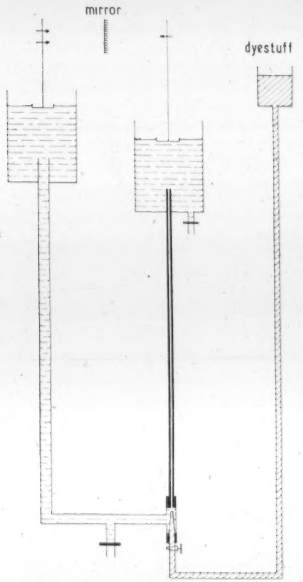


Fig. 6. Apparatus for the simultaneous measurement of velocity and viscosity ratios.

mixtures of the same density (1 cm^3 of Teepol was added per 100 g of beads to stabilize the sols). An indicator fluid consisting of a blue dye in a glycerol–water mixture of the same density as the suspensions was employed for the measurement of maximum velocities.

Each experiment involved the measurement of the relative viscosity of the sol and the corresponding ratio of its mean to its maximum velocity of flow in the capillary. Values of the former ratio were obtained by timing the flow of a fixed volume of a given suspension, and of an equal volume of the suspension fluid alone, between fixed levels indicated by coincidences of the two cross-wires, 0·5 cm apart, and carried by the one float, with the image of the cross-wire carried by the other float, the flow taking place under a mean pressure head of 8 cm of sol. The maximum velocity was obtained by timing the leading tip of a stream of dyestuff admitted to the sol under a pressure just sufficient for it to emerge from the jet, between fixed marks on the capillary tube. The mean velocity values, and thence the values of the velocity ratio were obtained by dividing the rate of flow of the sol by the sectional area of the tube.

Relative measurements on solvent and sol were made in rapid succession to minimize the temperature changes; the temperature was held constant throughout the whole range of experiments to 20 $\pm$ 2° c. Since relative viscosity values

have small temperature coefficients, this precaution was adequate without the complication of a thermostatic bath.

Precautions were taken to ensure the stability of the sols throughout the series of experiments. These included the arrangement of the vertical limb tubes with their open ends central in the reservoirs, the arrangement of experiments to occupy minimum time, and the use both of very large volumes of sol and of velocity and viscosity ratios obtained from quantities measured almost simultaneously.

Three sets of experiments as described above were carried out. In each set, experiments were made to determine corresponding values of the viscosity and velocity ratios for each of a series of sol concentrations from zero, increasing by 5% concentration steps, to 40%. Sets differed only with regard to the internal diameter of the capillary used; in the last two sets capillaries of 0·186 cm and 0·280 cm replaced the capillary of 0·139 cm diameter previously mentioned.

4.2. *Experimental results*

The results obtained are indicated in Fig. 7 in which curves B, C and D relate to experiments carried out with

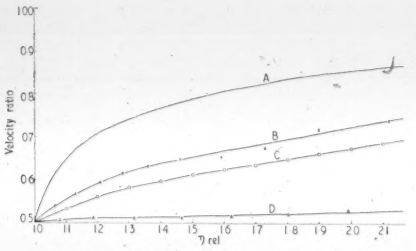


Fig. 7. Variation of velocity ratio with relative viscosity using tubes of diameter B, 0·139 cm; C, 0·186 cm; and D, 0·280 cm. (Curve A is theoretical).

capillary tubes whose internal diameters are 0·139 cm, 0·186 cm, and 0·280 cm respectively. Curve A is plotted from Eqn (13) of §1 of this paper.

The experimental curves show that the velocity ratio in all cases increases from a minimum value of 0·5 at zero concentration to a maximum value which is always less than unity. This implies a velocity profile which changes progressively with relative viscosity from the parabolic type characteristic of an approximately Newtonian fluid at a relative viscosity of unity to a flattened paraboliform profile in which the central flattened region extends nearer to the tube wall as relative viscosity increases. Further, the curves show that profile changes occur most rapidly in the lower part of the relative viscosity range.

In carrying out the experiments, curve B was obtained first. Although it has the same general form as the theoretical curve A, it was found not to be coincident with it. The discrepancy between these two curves can however be accounted for qualitatively in every detail in terms of the entrance effect discussed in §1.2.8. (The theoretical curve, it will be remembered, is obtained from an equation in the derivation of which the entrance effect is neglected and which therefore would be expected to apply quantitatively only when very long capillary tubes are used. The tubes used in the present experiments were relatively short.) As explained there the entrance effect would be expected to have greater significance, other conditions remaining the same, in wider bore tubes. It was for this reason that the experiments yielding curves C and D were carried out. As the results show, discrepancies between theoretical and experimental values diminish with a decrease in tube radius, that is as the entrance effect is reduced.

Attempts were made to use concentrations in excess of 40% but these led to considerable difficulties as a result of the continual blockage of the finest tube used, presumably through thrombosis.

5. Summary and conclusions

5.1. *Summary of the results*

The mechanism on which the present work is based, depends upon the existence of least action forces which arise during the shearing of a sol, and operate on the particles of the disperse phase, so as in general to displace them from the streamlines on which they initially move, to other streamlines where the velocity gradient is lower. The operation of these forces results in the development, within a sol under shear, of a concentration pattern, characterized by an absence of particles in the highest shear rate region of the system, and a concentration of particles where the lowest rates of shear prevail. (Under appropriate conditions such patterns can be made apparent (Starkey and James 1956).)

By relating mathematically various parameters associated with the concentration pattern characteristic of flow through a capillary tube, the following conclusions were reached:

(1) Viscosity shear dependence is a characteristic of all suspensoid sols and observable increases in apparent viscosity must be expected to occur if the flow rate is reduced sufficiently.

(2) This shear dependence of viscosity is due to changes in the parameters defining the concentration pattern; these can be accounted for in terms of changes in the magnitudes of least action forces with flow rate.

(3) The mean concentration of a suspensoid sol flowing through a capillary tube is always less than the concentration in the reservoir from which it is derived.

(4) The mean concentration in an element of a capillary taken over the whole cross section, varies along the tube from a maximum value (equal to the reservoir concentration) at the tube entrance, to a minimum value at some point nearer to the exit.

(5) The mean overall concentration in a capillary is shear dependent; it lies in the case of flow through a long capillary tube between theoretical limits of one half of the reservoir concentration and the reservoir concentration itself.

(6) The velocity profile in a tube (as defined by the ratio of the mean to the maximum velocity) is shear dependent. This ratio also lies between theoretical limits of one half and one.

(7) The ratio of the mean velocity to the maximum velocity, is, for a sol whose particles show no appreciable Brownian movement, and in a long tube, related with the relative viscosity, η_{rel} measured under the same conditions, by

$$\frac{v_m}{v_p} = \frac{\sqrt{\eta_{rel}} + \sqrt{(\eta_{rel} - 1)}}{2\sqrt{\eta_{rel}}}$$

where v_m and v_p are the mean and maximum velocities respectively.

(8) Under the conditions stipulated in (7) above, the ratio of mean to maximum velocity is equal to the ratio of mean tube concentration to reservoir concentration.

(9) The maximum concentration of the sol flowing in a tube is always greater than the concentration in the reservoir.

(10) The value of the concentration in the axial region varies systematically from point to point along the length of the tube. At the entrance it is equal both to the mean concentration and to the reservoir concentration, but increases to a maximum value as the exit end is approached.

(11) The hyperconcentration which develops in the axial region of a tube can under appropriate conditions reach extremely high values. It is presumably the cause of the formation of a thrombus in coronary thrombosis.

The experimental work described in the preceding sections of this paper was designed for the investigation of the theoretical conclusions numbered (1), (3), (5), (6) and (7) above. The results obtained support these conclusions.

5.2. *Some consequences of the results*

In the light of these results the method of deriving viscosity–concentration relationships initiated by Einstein (1906) and subsequently developed by various authors in attempts to account for a variety of flow phenomena in suspensoid sols, seems to be fundamentally unsound in the respect that it is based on the assumption that the random distribution of the particles throughout a suspensoid sol at rest, holds also under shear. The present results suggest that in a stable suspensoid sol under shear a concentration pattern exists; this is not consistent with the assumption of random particle distribution.

The use of the assumption of random particle distribution within a sol can lead at best therefore only to approximate results. For sols, and particularly for molecularly dispersed sols, undergoing shear in relatively short and wide tubes, there is, however, no reason to expect the approximation to be other than close. The measure of agreement between the viscosity–concentration relationship derived by Einstein and experimental results obtained under the conditions just mentioned, has, however, apparently encouraged the making of the further assumptions mentioned previously regarding the parts played by the various characteristics of the particles of the disperse phases, as a basis for explaining the shear dependence of viscosity and other flow phenomena. Whilst these further assumptions may or may not have some relevance in particular cases they cannot completely account for these phenomena.

Acknowledgments

We gratefully acknowledge our indebtedness to the Monmouthshire Local Education Authority for providing facilities for the experimental work described in this paper and to the Department of Scientific and Industrial Research for a special research grant to assist with this work. We are indebted also to Messrs. British Nylon Spinners Ltd., to Messrs. BX. Plastics Ltd., and to other firms for sundry gifts of equipment and materials.

We also acknowledge our indebtedness to Dr. G. W. Roderick, Mr. H. S. Carter, Mr. D. Bridges, and Mr. G. Evans who are all members of the staff of the College, for their help with various aspects of the experimental work.

References

EINSTEIN, A., 1906, *Ann. Phys. Lpz.*, **19**, 289.

HARLAND, W. G., 1955, *Shirley Inst. Mem.*, **28**, 111, *J. Text. Inst.*, **46**, T472.

HIGGINBOTHAM, G. H., OLIVER, D. R., and WARD, S. G. 1958, *Brit. J. Appl. Phys.*, **7**, 372.

IMMERGUT, E. H., and EIRICH, F. R., 1953, *Industr. Engng Chem.*, **45**, 2500.

RICHARDSON, E. G., and TYLER, E., 1933, *Proc. Phys. Soc.*, **45**, 142.

SCOTT BLAIR, G. W., 1958, *Rheol. Acta*, Nummer 2/3, Seite 123–126.

STARKEY, T. V., 1955, *Brit. J. Appl. Phys.*, **6**, 34.

—— 1956, *Brit. J. Appl. Phys.*, **7**, 52.

STARKEY, T. V., and JAMES, R. E., 1956, *Nature (London)*, **178**, 207.

VEJLENS, G., 1938, *Acta. Pathol.*, Suppl. **23**.

WORSNOP, B. L., and FLINT, H. T., 1931, *Advanced Practical Physics*, 3rd edn (London: Methuen), p. 150.

Diamond knife ultra microtomy of metals and the structure of microtomed sections

by R. PHILLIPS, B.Sc., A.Inst.P., Aeon Laboratories, W. Wittley and Co. Ltd., Beech Hill, Ridgemead Road, Egham, Surrey

MS. received 18*th January* 1961, *in revised form* 26*th May* 1961

Abstract

The parameters of microtome sectioning are defined, and the use of the diamond knife ultra microtome for metallurgical studies is described. The effect of cutting stresses on the formation of sections is discussed. It is reasoned that vibration of parts of the cutting system would only be excited by the first impact of the specimen or the knife and not by the forces acting on them throughout cutting. Vibration is shown to produce smooth variations in section thickness, the thickness being constant all along any normal to the direction of cutting, and varying along the direction of cutting. Knife marks are shown to be due to variations in cutting stress associated with knife edge defects; these produce variations in the thickness and structure of the section, even when the mechanism proposed by Farrant cannot apply. Some folds and wrinkles are explained as being due to variation in cutting stress due to variations in knife edge quality. All the above defects can be avoided by a proper choice of experimental conditions, but it is shown that there is a shear stress imposed on the section during cutting which in the majority of cases permanently deforms it, producing a surface structure characteristic of the material being cut.

Introduction

ULTRA-THIN sectioning is well established as a technique for the preparation of embedded biological tissues for transmission electron microscopy. A number of workers (Fernandez-Moran 1956a, Reimer 1959a, b, Menter 1958, Haansbra 1955, Phillips 1961) have extended the application of the technique to cover pure metals and alloys, and other hard materials. It has been widely felt that such sections do not provide useful metallurgical and physical information about the sample from which they were obtained because of the stresses which are present during the cutting process. The exact nature and effects of these stresses have not, however, been investigated. The purpose of this paper is to report the results of the initial work in a long-term study of the processes involved in cutting sections, with particular emphasis on the processes involved in cutting metals, and the information this provides about the sectioning of other materials. This type of study is relevant to two broader fields of applied science—the machining of metals and the application of microtomy to electron-metallography.

More detailed descriptions of the application of the ultra microtome to both these subjects will be described in later papers. Here the main consideration is the effect of the cutting processes on the structure of the section obtained with the ultra microtome.

The ultra microtome

The ultra microtome (Fernandez-Moran 1956b, 1957, Porter and Blum 1953, Hellstrom 1960, Huxley 1959, Farrant and Powell 1956, Sitte 1955) is a machine which takes a specimen past a keen cutting edge at intervals while providing a very small advance of the specimen toward the knife edge during each interval, so that, once the specimen meets the knife, a thin section is removed from the specimen each time it is brought past the cutting edge.

A reservoir of liquid is maintained behind the cutting edge, and the sections may subsequently be collected from the surface of this liquid. In all the work described in this paper, the cutting edge is a specially prepared diamond (Fernandez-Moran 1956a).

Definition of terms and parameters

Figure 1 shows a schematic representation of the relation of the specimen to the knife during microtomy. It also

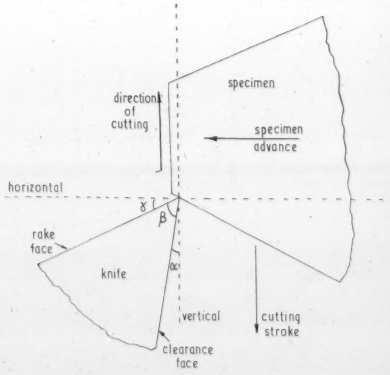


Fig. 1. Schematic representation of cutting system in microtome sectioning and definition of parameters.

defines the terms used in this paper to describe the various cutting parameters which are as follows:

The plane of cutting: The plane defined by the knife edge and the direction of cutting.

Clearance angle α: The angle between the front face of the knife and the plane of cutting.

Knife angle β: The angle between the two polished faces of the knife (the rake and clearance faces).

Rake angle γ: The angle between the upper face of the knife and the normal to the plane of cutting.

The preparation of specimens for sectioning

In order to avoid unnecessary knife wear, and to minimize the stresses set up in the instrument, it is desirable to cut sections of the smallest area compatible with reasonable ease of handling and with the extraction of the required information from them. The specimen grips require a cylinder about 0·5 cm in diameter; therefore a tip has to be formed at the end of the cylinder so that a much reduced dimension is presented to the knife edge. A typical section area for the softer metals is 0·1 mm × 0·1 mm, though for very hard metals it may be much smaller than this, and for biological specimens the linear dimension may be as high as 1 mm. Because sections are more easily handled and may be studied as a series if they are produced edge to edge in a straight line, biological workers have trimmed right pyramidal tips on their specimens. The sections then have parallel sides which facilitates the formation of the desired 'ribbon'. Biological specimens are trimmed with a razor blade, but metals and hard materials require a special technique. A jig has been devised which enables the four faces of a right pyramid with a semi-angle of 50° (to provide adequate rigidity without a too rapid increase in area during sectioning) to be prepared by normal metallographic grinding down to 4/0 grid emery paper (i.e. each face of the pyramid is ground on successively finer grades, the direction of grinding being changed by 90° at each change of paper). Comparison of the arcing of the maxima in selected area diffraction patterns from sections cut off a tip of copper prepared in various ways (fine filing, coarse grinding), has shown that this procedure is necessary if the volume which will subsequently be sectioned is not to be severely worked in the tip preparation.

Serial sections are not so important in studies of metals, and in the case of very hard material it may be advantageous to dispense with the easier handling allowed by ribbon formation in order to reduce the impact when the specimen first meets the knife. Thus it may be better to prepare a three-sided tip, and orient it so that the knife meets one of the edges common to two of the sides of the tip. A triangular section would then be cut with the cutting direction running from apex to base.

Setting up the microtome

For minimum distortion of the section during cutting, and for maximum resistance of the knife edge to fracture and wear, the clearance angle must be a minimum, thus making the rake angle a maximum for a given knife angle. Even using simple ancillary optical equipment, the best accuracy that can be obtained in the setting of the clearance angle is to about 0·5°. As the clearance angle must be positive in order to avoid compressing the specimen after cutting, it seems best, therefore, to aim at a clearance angle setting of 0·5° to 1°.

Next, the reservoir behind the knife should be filled with a suitable liquid so that the meniscus at the cutting edge is horizontal and all the edge is wetted. Distilled water has a good combination of physical properties—relatively high surface tension, low viscosity and low rate of evaporation. It therefore provides easy transport of the sections from the knife edge as other sections are produced, allows the sections to float so that they are easy to collect, and avoids the need for frequent adjustments to the meniscus. Distilled water is also clean, but otherwise its chemical properties are less attractive than the physical. Sections of metals which oxidize very readily convert completely to oxide when on water, and in this case mechanically less desirable fluids, e.g. xylol, have to be used.

The main precaution in commencing sectioning (apart from the obvious one of adequately clamping all adjustments and vices) is to ensure that the first cut is of the same order of thickness as that required in the sections for the electron microscope. The stereoscopic binocular microscopes universally used on microtomes, however, have such a low numerical aperture that they cannot resolve less than about 5 microns under typical illuminating conditions. Thus it is not possible to approach the knife edge sensitively with visual control. The 5-micron gap must be closed with fine adjustments, with the employment of the microtome advance mechanism in such a way that the first cut removes less than 1000 Å from the specimen. If this is not done, the portion of the knife in use will not be at its best for cutting thin sections, and in the case of hard materials will probably be spoilt before any useful sections are obtained.

Features arising in sections due to the cutting conditions

In addition to features corresponding to the structure of the specimen from which they were cut, sections may show features due to (i) instrumental defects, (ii) the stresses which exist in the section during cutting, (iii) the local variations of the cutting stresses due to the instrumental defects.

(i) *Features due to instrumental defects*

Instrumental defects may be subdivided into knife edge defects and unwanted movements of the specimen with respect to the knife edge.

(a) Knife edge defects

In general a chip in the knife edge means that the rake angle is less, that the cutting edge is farther back and the clearance angle may locally be negative. The decrease in rake angle produces greater strain in the section, which results in greater section thickness or local fractures along the track, thus producing contrast along the track of a defect (knife mark). Knife edge defects also leave grooves on the cut surface. Because the local clearance angle may be negative, local compression and surface roughening may also occur along the grooves, and this affects the contrast in the next section.

(b) Extraneous movements of the specimen with respect to the knife edge

The unwanted movements of the specimen with respect to the knife edge are any which are in addition to the intended displacement of the specimen towards the knife (specimen advance) and the cutting stroke. The unwanted movements may be random or periodic and may be resolved into three components, parallel to the knife edge, parallel to the direction of specimen advance and parallel to the direction of cutting.

Random movements. Farrant and Powell (1956) have pointed out that a displacement of the specimen along the knife edge between cuts produces variations in section thickness across the tracks of defects in the cutting edge (Fig. 2(*a*)). If this type of displacement can be avoided, the sections will be of much more uniform thickness along the

direction of viewing, even if the cutting edge contains defects (Fig. 2(*b*)), although there will still be contrast on the tracks of the defects as explained above.

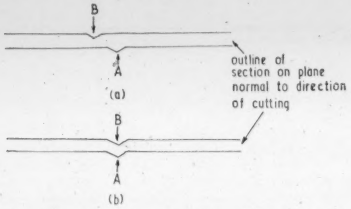


Fig. 2. (*a*) Variations in section thickness due to excursion of specimen along defective knife edge between cuts. (*b*) Approximately constant thickness of section cut by defective knife edge, despite lateral excursions of the specimen between cuts.

Figure 3 shows coarse parallel lines of different widths, in both dark and light contrast. These run in the direction of cutting and are due to knife edge defects. Random movements in the direction of specimen advance produce

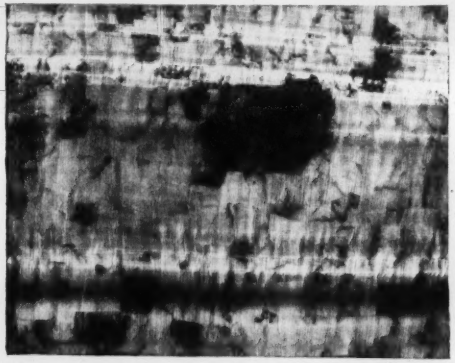


Fig. 3. Knife marks on an aluminium section. ×9000 (in all the micrographs the cutting direction is approximately horizontal).

wide irregular variations in thickness along a section or from section to section. Random movements in the direction of cutting produce variations in cutting speed. The final effect of this is as yet unknown, but is unlikely to be very significant.

Periodic movements (*vibrations*). In line with engineering practice, which has also been followed in the definition of angles, the term chatter will be used to describe relative vibration of the specimen and knife, and its effect.

Vibration in which the amplitude is parallel to the knife edge is unlikely to be excited during sectioning; that in which the amplitude is parallel to the direction of cutting is one of the more probable practical cases, but as in the case of the equivalent random movement is unlikely to have an important effect. Periodic movement along the direction of specimen advance produces regular variations in the thickness of section (unless the vibrations in successive cuts are exactly in phase) and a regularly undulating cut surface on the specimen. These effects are not to be confused with the fine scale line structure later interpreted as slip lines due to cutting stresses. The peaks (and troughs) of the undulations will, of course, be normal to the direction of cutting, and be continuous across the section and cut area. These features are aids in deducing when 'chatter' has been excited. Simple formulation of the problem shows that the forces involved in the cutting process are unidirectional and non-periodic. Calculation shows that they are insufficient to produce a significant distortion of the system (Brown and Phillips 1961). Thus the stresses involved during the equilibrium cutting of metal are not of the correct type, and are insufficient in magnitude for producing chatter. It is thought therefore that chatter is vibration set up by the first impact of the specimen with the knife. The fact that lower speeds are less conducive to chatter is consistent with this interpretation, which has also been presented by Gettner and Ornstein (1956). The location of the vibrating member has not yet been found.

(ii) *Features due to cutting stresses*

Sections of pure metals are often characterized by fine lines (Fig. 4) running approximately normal to the direction

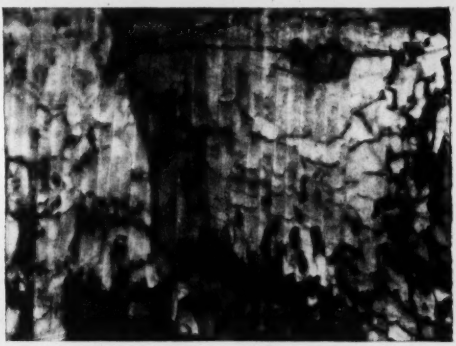


Fig. 4. Fine lines on an aluminium section. ×100000.

of cutting. These lines are associated with the surface topography of the sections as is proved by the following experimental observations:

(*a*) There are Fresnel fringes associated with the lines in out-of-focus micrographs (i.e. Fig. 4 is slightly underfocused) suggesting that there is a sharp discontinuity in section thickness at each line.

(*b*) Replicas of the top side of sections show a structure on the same scale as the fine lines. Fig. 5 is a replica prepared by Bradley (1959) from a section similar to that photographed in transmission in Fig. 6. Replicas from the bottom of the sections and from the cut tip of the specimen do not appear to show structure in this direction.

(*c*) Sections which have been embedded and resectioned in a plane parallel to the original direction and normal to the original plane of cutting usually show a pronounced structure on the side corresponding to the top of the original section. Sufficiently thin sections of a section have not yet been obtained however, for a detailed analysis of the structure to be made.

It must be emphasized that these lines are very sharp and

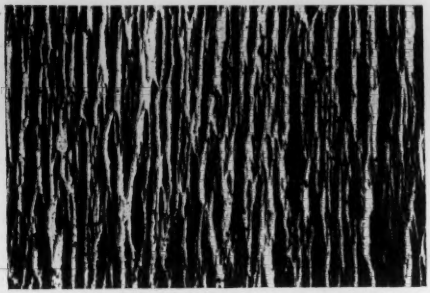


Fig. 5. Platinum–carbon replica of the top surface of an aluminium section. ×8000.

Fig. 6. Aluminium section cut from same specimen, very close to and under the same condition as that replicated in Fig. 5. ×8000.

may be quite short and very finely spaced, in contrast to chatter markings which are broad and extend right across the section.

Workers in the field of the physics of machining have shown that a plane of high shear stress exists from the tool tip to the outer surface of the chip (Ernst 1938, Merchant 1950). Fig. 7 is reproduced from Merchant's paper. The

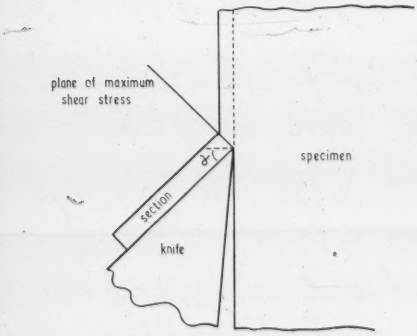


Fig. 7. Schematic representation of a section (in which plastic deformation occurs (after Merchant (1950)).

shear stress arises because the material being cut from the block has to be transferred from the plane of cutting to the rake face of the knife, and is greater the smaller the rake angle. It is suggested that the fine line structure of Fig. 4 represents slip lines formed in the relief or partial relief of the shear stress. It is an immediate consequence of this deformation during cutting that the section is usually thicker than the amount of specimen advance, and that the material is shorter in the direction of cutting in the section than it was before cutting in the specimen. The amount of thickening and shortening involved increases as the rake angle decreases.

Electron diffraction shows that the change in orientation of face-centred cubic metals on cutting is often that which would be expected on the shear theory of section formation. There is usually, however, a second orientation present which is not accounted for on the simple theory and which may indicate either that more careful experimentation is required to achieve the initial assumptions of the theory, or that the theory itself requires elaboration.

It may be mentioned here that the effect of the shear stress on the structure of the section depends on the material being sectioned. The shear stress, as has been explained, produces slip lines on sections of ductile metal. In other cases it produces homogeneous plastic shear, fracture and/or purely elastic distortion. The conditions aimed at in the sectioning of biological blocks is that the distortion shall be purely elastic and be completely removed when the section is free of stress at the end of the cut. The fact that most biological sections are shorter than the face from which they were cut indicates that this is rarely the case in practice. It would further seem that compression is a slightly misleading term for this shortening which plasticity theory shows to be mainly due to shear strain. Details of the types of sections formed from a wide range of materials and the mechanisms involved will shortly be published at length.

The evidence submitted here, together with that presented by others, suggests that the very fine scale surface structures revealed on methacrylate sections by low beam techniques in the electron microscope (Morgan, Moore and Rose 1956) and heavy metal shadowing (Williams and Kallman 1955) are the result of the deformation of the impregnated tissue in the section under the stress which acts in it during cutting. The comments of the above workers seem to be in agreement with this, but discussion has since been confused by authors who refer to these structures as a form of chatter.

(iii) *Effects due to local variations of cutting stresses*

Optical examination of metal sections almost invariably shows them to be rumpled. This effect is also sometimes found in methacrylate and araldite. In severe cases, material is folded right back on itself, and a transmission electron micrograph of a badly folded section is shown in Fig. 8. The processes involved in forming rumples and folds are not fully understood; their formation is minimized by the use of a smooth, clean, lubricated knife face, sharp cutting edge, large rake angle and correctly adjusted liquid meniscus. However, much more folding is encountered in copper than in aluminium, for example, and its incidence seems to vary with thickness of cut. The details of the form of the folds and rumples are controlled to some extent by the degree and distribution of knife marks.

In a simple picture of folding, the part of the section already cut may be held up on the knife face by other sections or the form of the liquid meniscus, so that a following part of the section has to bow out over the first part.

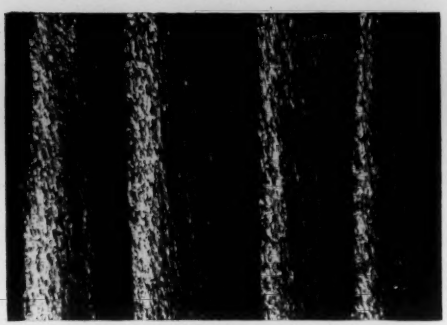


Fig. 8. Badly folded section of copper. ×8000.

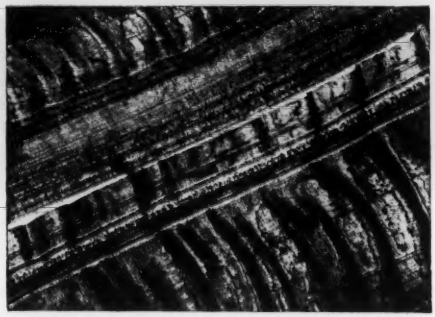


Fig. 9. Fold bounded by knife marks in an aluminium section. ×4000.

Figure 9 shows a section in which rows of folds or wrinkles are bounded by knife marks. The rake angle is less at knife marks, and produces a greater shear strain which in turn results in a specimen which is thicker and shorter along the band of the knife mark, than where the cutting edge was more perfect. The only way in which the differences in length can be accommodated without tearing is by the buckling of the longer portions.

Conclusions

Thin sections obtained with an ultra microtome may show a number of features which are purely due to the cutting stresses and the processes they excite in the specimen. These features and their origins may be analysed as follows:

(i) Instrumental vibrations—these may produce a wavy section, or one with a periodic variation in thickness. Tests have been suggested which will indicate when vibrations are occurring. The amplitude of the vibrations in a given specimen will be reduced by decreasing the width of the area cut, decreasing the section thickness, and decreasing the cutting speed.

(ii) Knife marks and tears—these are due to defects of, and on, the cutting edge, which when combined with the effects of lateral excursion of the specimen produce variation in thickness across the section or tears. Even in the absence of lateral excursions, the specimen has contrast along knife marks in transmission micrographs due to a different structure which arises from the increased cutting stresses associated with the defect.

(iii) Folds and wrinkles—these arise because of resistance to the free passage of the section from the knife face and from variation in the quality of the knife edge.

(iv) Surface structure of the section—this arises because of fracture or deformation due to the shear stress which operates during thin sectioning.

As expressed here, these conclusions apply to biological sections and sections of plastics as well as to metal sections, the exact incidence of the different features depending on the mechanical properties of the material being sectioned.

Acknowledgments

The author is grateful to A. W. Agar, Head of Aeon Laboratories for helpful discussions and for permission to publish this paper. The research reported in this document has been made possible through the support and sponsorship of the U.S. Department of the Army through its European Research Office.

References

BRADLEY, D. E., 1959, *Brit. J. Appl. Phys.*, **10,** 198.
BROWN, A. J., and PHILLIPS, R., 1961, *Aeon Laboratories Report*, FTR. 2.
ERNST, H., 1938, *Machining of Metals*, pp. 1–78 (Cleveland, Ohio: ASM).
FARRANT, J. L., and POWELL, S. E., 1956, *Proc. 1st Regional Conf. Elec. Mic., Japan*, 146.
FERNANDEZ-MORAN, H., 1956a, *J. Biophys. Biochem. Cytol.*, **2** (Suppl.), 29.
—— 1956b, *Industr. Diam. Rev.*, **16,** 128.
—— 1957, *Mikroskopie*, **12,** 81.
GETTNER, M. E., and ORNSTEIN, L., 1956, *Physical Techniques in Biological Research III* (London: Academic Press), p. 658.
HAANSTRA, H. B., 1955, *Philips Tech. Rev.*, **17,** 161.
HELLSTROM, B., 1960, *Sci. Tools*, **7,** 10.
HUXLEY, A. F., 1959, *Ultra Microtome* (Camb. Inst. Co. Technical Brochure).
MENTER, J. W., 1958, *Adv. Phys.*, **7,** 299.
MERCHANT, M. E., 1950, *Machining, Theory and Practice*, 5–44 (Cleveland, Ohio: ASM).
MORGAN, C., MOORE, D. H., and ROSE, H. M., 1956, *J. Biophys. Biochem. Cytol.*, **2** (Suppl.), 21.
PHILLIPS, V. A., 1961, *European Reg. Conf. Elec. Mic., Delft*, 1960, **1,** 485.
PORTER, K. R., and BLUM, J., 1953, *Anat. Rec.*, **117,** 685.
REIMER, L., 1959a, *Metallkunde*, **50,** 37.
—— 1959b, *Naturwissenschaften*, **46,** 68.
SITTE, H., 1955, *Mikroskopie*, **10,** 365.
WILLIAMS, R. C., and KALLMAN, F., 1955, *J. Biophys. Biochem. Cytol.*, **1,** 301.

Capstan equation for strings with rigidity

by I. M. STUART, Commonwealth Scientific and Industrial Research Organization, Wool Research Laboratories, Division of Textile Physics, Ryde, N.S.W., Australia

MS. received 13*th March* 1961

Abstract

The effect of string rigidity on the capstan equation is considered. Two cases which cover the most likely situations are distinguished—contact of the string with the peg at a point, and contact over an arc. The analysis of these situations is facilitated by an approximation derived for the curvature for the leads of string on either side of the peg. If the friction is governed by Amontons' law the correction of the capstan equation for string rigidity will usually be small, but in the case of arc contact and load-dependent coefficient of friction, the correction could be important.

1. Introduction

When a taut string slides round a peg, friction causes the tension to vary along the string. If the string is perfectly flexible, the friction limiting and described by Amontons' law, the tensions T_1, T_2 on the tauter and slacker sides respectively are related to Φ, the angular displacement of the string, and to μ, the coefficient of friction, by the capstan equation

$$\frac{T_1}{T_2} = e^{\mu\Phi}. \qquad (1)$$

If the string has rigidity it will bow out, so that the tangents to the string at the ends of its arc of contact with the peg will be inclined to the lines of action of the forces pulling on the string ends. Thus the angular displacement of the string while it is in contact with the peg will be less than the total angular displacement of the string between the points of application of the holding forces. This point was made by M. W. Pascoe (see Howell, Mieskes and Tabor 1959, p. 50), who suggested that this bowing could cause a reduction in the tension drop from the tauter to the slacker side of the peg—assuming the capstan equation (Eqn 1) holds on this arc of contact.

In this paper we discuss two simple cases which can arise—point contact and arc contact. Point contact may arise if the angular displacement of the string is small, the tensions low, or the radius of curvature of the peg small. The conditions for such point contact are derived, and the appropriate 'capstan' equation is established.

If the contact is not at a point then it should be over an arc. In this case, in addition to distributed reaction forces between the string and the peg (T/ρ per unit length of string for a circular cylindrical peg of radius ρ) there are concentrated reaction forces acting at the ends of the arc of contact. Within the arc of contact the shear force in the string is governed by the peg radius of curvature (in case of a circular peg section, the shear force in string is zero in the arc of contact), whilst outside the arc of contact the shear force in the string balances the component of the holding force normal to the string. Thus a concentrated reaction at an end of the arc of contact balances the difference between the shear forces just outside and just inside the arc. The friction forces associated with these concentrated reactions make up the reduction in tension drop caused by bowing to the first order, so that corrections for rigidity to the capstan equation are of second order, and usually negligible.

2. Approximation of curvature in leads

In the subsequent analysis we use an approximation for the curvature in the leads of string to the peg:

$$\frac{d\theta}{ds} = 2\left(\frac{P_0}{B}\right)^{1/2} \sin \tfrac{1}{2}\theta$$

where P_0 is the magnitude of the holding force, B the rigidity modulus of the string and $\pi - \theta$ the angle between the tangent to the string at the current point and the holding force $\mathbf{P}_0$.

The error in this approximation is small when (i) the manner of holding is restricted, (ii) θ is less than π, (iii) the length x of the projection of the lead of string on to the line of action of the holding force is large compared with a length c, where $c = (B/P_0)^{1/2}$, so that $2\exp(-2x/c) \ll 1$. The restriction on the manner of holding will be discussed below.

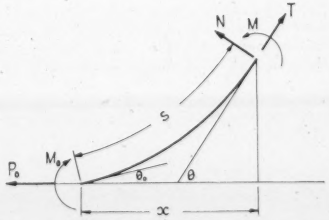


Fig. 1. Forces and moments on a lead of string.

With s the curvilinear distance along the string from the point of holding, $s = 0$, $\pi - \theta$ the angle between the string and the force $\mathbf{P}_0$ (see Fig. 1) we define c as above, $\sigma = s/c$, $\phi = \frac{1}{2}\theta$. Then the curvature in the string is given by

$$\frac{d\theta}{ds} = 2\left(\frac{P_0}{B}\right)^{1/2} \frac{d\phi}{d\sigma},$$

and the bending moment

$$M = B\frac{d\theta}{ds} = 2(BP_0)^{1/2}\frac{d\phi}{d\sigma}.$$

The end conditions are, at $\sigma = 0$, $\phi = \phi_0$ and $d\phi/d\sigma = (d\phi/d\sigma)_0$.

Balancing the forces acting on the interval of string $(0, s)$ normal to the tangent at the current end, we obtain for the shear force $N = -P_0 \sin 2\phi$.

The equation derived from the relationship between the shear force and the bending moment $N = -dM/ds$ reduces to

$$\frac{d^2\phi}{d\sigma^2} - \sin\phi\cos\phi = 0.$$

This equation can be written in either of the forms

$$\frac{d}{d\sigma}\left(\frac{d\phi}{d\sigma} \pm \sin\phi\right) \mp \cos\phi\left(\frac{d\phi}{d\sigma} \pm \sin\phi\right) = 0,$$

which integrate to give the two equations

$$\frac{d\phi}{d\sigma} \pm \sin\phi = \left\{\left(\frac{d\phi}{d\sigma}\right)_0 \pm \sin\phi_0\right\} \exp\left(\pm\int_0^\sigma \cos\phi d\sigma\right).$$

By a simple manipulation we then obtain

$$\frac{d\phi}{d\sigma} = \sin\phi\left(1 + \frac{2\alpha e^{-\beta}}{1 - \alpha e^{-\beta}}\right) \tag{2}$$

where

$$\alpha = \left\{\left(\frac{d\phi}{d\sigma}\right)_0 - \sin\phi_0\right\} \Big/ \left\{\left(\frac{d\phi}{d\sigma}\right)_0 + \sin\phi_0\right\}, \beta = 2\int_0^\sigma \cos\phi d\sigma.$$

Eqn (2) is exact within the beam approximation hypothesis. The approximation to be used is obtained by neglecting the second term in the parentheses:

$$\frac{d\phi}{d\sigma} \simeq \sin\phi \text{ if } \frac{2\alpha e^{-\beta}}{1 - \alpha e^{-\beta}} \ll 1.$$

If $(d\phi/d\sigma)_0$ and $\sin\phi_0$ are of the same sign, $|\alpha| < 1$. The extremes of this range represent the well-known end conditions 'clamped' and 'freely hinged'. When the end $s = 0$ is clamped $\phi = 0$, so $\alpha = 1$; when freely hinged, $(d\phi/d\sigma)_0 = 0$ and $\alpha = -1$.

If ϕ is limited to be in the range $-\frac{1}{2}\pi < \phi < \frac{1}{2}\pi$,

$$\cos\phi > \cos 2\phi \text{ and } \beta > 2\int_0^\sigma \cos 2\phi d\sigma = \frac{2x}{c},$$

where x is the length of the projection of the string on to the line of action of the force $\mathbf{P}_0$, as indicated in Fig. 1.

So, if $\sin\phi_0$ and $(d\phi/d\sigma)_0$ are of the same sign, and $x \gg c$ and positive, so that $2\exp(-2x/c) \ll 1$, $d\phi/d\sigma \simeq \sin\phi$, with a fractional error of the order of $2\exp(-2x/c)$.

Restriction on manner of holding

If x is to be positive $|\theta_0|$ would normally be less than $\frac{1}{2}\pi$, and, if this is so, then $\sin\phi_0$ and $(d\phi/d\sigma)_0$ will be of the same sign if both the string and the centre of curvature of the string lie on the same side of line of action of the holding force $\mathbf{P}_0$.

In what follows it will be assumed that such end conditions apply, and that $x \gg 2c$, and we will write

$$\frac{d\theta}{ds} = 2\left(\frac{P_0}{B}\right)^{1/2} \sin\tfrac{1}{2}\theta. \tag{3}$$

3. String in point contact with the peg

The string is assumed to be in point contact with the peg at P, where the peg radius of curvature is ρ (see Fig. 2). At P the string is inclined at angles θ_1, θ_2 to the lines of action of the holding forces T_1, T_2, so that the angle between the directions of pulling on the string,

$$\Phi = \theta_1 + \theta_2. \tag{4}$$

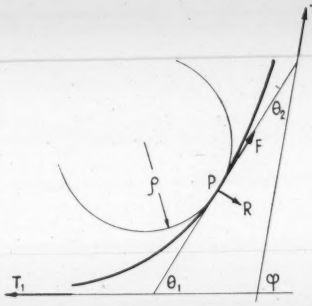


Fig. 2. Forces operating on string in point contact with a peg radius ρ.

Condition for point contact

Assuming conditions in the leads of string to the peg to be suitable for the application of the approximation derived in §2, the curvature in the string at P, κ is given by either of the two formulae:

$$\kappa_P = 2\left(\frac{B}{T_1}\right)^{1/2} \sin\tfrac{1}{2}\theta_1; \ \kappa_P = 2\left(\frac{B}{T_2}\right)^{1/2} \sin\tfrac{1}{2}\theta_2.$$

The condition for point contact is that, at P, the curvature in the string be less than or equal to the curvature of the peg. That is

$$\kappa_P \leqslant \frac{1}{\rho}.$$

Whence

$$2T_1^{1/2} \sin\tfrac{1}{2}\theta_1 = 2T_2^{1/2} \sin\tfrac{1}{2}\theta_2 \tag{5}$$

$$\leqslant \left(\frac{B}{\rho^2}\right)^{1/2}. \tag{6}$$

Using Eqns (4), (5), (6) in the identity

$$\sin^2(\tfrac{1}{2}\theta_1 + \tfrac{1}{2}\theta_2) = \sin^2\tfrac{1}{2}\theta_1 + \sin^2\tfrac{1}{2}\theta_2 + 2\sin\tfrac{1}{2}\theta_1 \sin\tfrac{1}{2}\theta_2 \cos(\tfrac{1}{2}\theta_1 + \tfrac{1}{2}\theta_2)$$

we obtain the condition for point contact in terms of T_1, T_2, ϕ, ρ and B.

$$4\sin^2\tfrac{1}{2}\Phi \leqslant \frac{B}{\rho^2 T_1}(1 + r^2 + 2r\cos\tfrac{1}{2}\Phi) \tag{7}$$

where

$$r^2 = T_1/T_2. \tag{8}$$

Normal reaction and friction forces

Equating forces parallel, then normal to the tangent at P, we obtain for the friction force

$$F = T_1\cos\theta_1 - T_2\cos\theta_2,$$

then, for the normal reaction at P

$$R = T_1\sin\theta_1 + T_2\sin\theta_2.$$

Using Eqn (5), these expressions reduce to

$$F = T_1 - T_2 \tag{9}$$

$$R = T_1^{1/2}T_2^{1/2}.\,2\sin\tfrac{1}{2}\Phi. \tag{10}$$

Thus, if the friction is described by Amontons' law, $F = \mu R$, the capstan equation is

$$T_1 - T_2 = \mu \cdot T_1^{1/2}T_2^{1/2} \cdot \sin\tfrac{1}{2}\Phi,$$

or

$$r - \frac{1}{r} = \mu \sin\tfrac{1}{2}\Phi.$$

whence

$$r = \mu \sin\tfrac{1}{2}\Phi + (1 + \mu^2 \sin^2\tfrac{1}{2}\Phi)^{1/2}. \tag{11}$$

To compare this formula with the capstan equation for flexible strings (which gives $r = e^{\frac{1}{2}\mu\Phi}$), we expand (11) in powers of Φ, to obtain

$$r = e^{\frac{1}{2}\mu\Phi}\left\{1 - \frac{\mu(1+\mu^2)}{3!}(\tfrac{1}{2}\Phi)^3 + \frac{\mu(1+\mu^2)(1+9\mu^2)}{5!}(\tfrac{1}{2}\Phi)^5 + \ldots\ldots\right\}. \tag{12}$$

Comments

The capstan equation for strings with rigidity in point contact is independent of the rigidity, or the radius of the peg.

The error involved in using the standard capstan equation is small for small angles. For $\mu < 1$, the error is less than 1% for $\Phi < \frac{1}{4}\pi$.

4. Contact over an arc

If the string does not touch the peg at a point, then it should come into contact over an arc. Contact at two points is a possibility but can be shown to be impossible for angular displacements of less than 180° or if the tension is to remain positive.

The string is assumed to be in arc contact with a peg of circular section, radius ρ on an arc AB. The tangents at A and B make angles θ_1, θ_2 with the lines of action of the holding forces T_1, T_2 respectively, as shown in Fig. 3. Then if the angle between these lines of action is Φ, the arc AB subtends an angle $\Phi - \theta_1 - \theta_2$ at the centre of the peg.

The curvature of the string on the arc AB is constant, and equals $1/\rho$. So (i) equating to $1/\rho$ the curvatures at A and B given by the approximation of §2 (Eqn 3) we obtain

$$2T_1^{1/2}\sin\tfrac{1}{2}\theta_1 = 2T_2^{1/2}\sin\tfrac{1}{2}\theta_2 = \left(\frac{B}{\rho_2}\right)^{1/2}, \tag{13}$$

(ii) the curvature and hence the bending moment in the string being constant on the arc AB, the shear force ($N = -dM/ds$) there is zero. So, in addition to a distributed loading on the arc there are concentrated normal reactions N_A and N_B at the ends of the arc, A and B, to balance the components of the forces T_1, T_2 normal to the tangents at A and B. Thus $N_A = T_1 \sin\theta_1$, $N_B = T_2 \sin\theta_2$.

The tensions just outside the arc AB are $T_1 \cos\theta_1$, $T_2 \cos\theta_2$. So, for the string sliding from B to A, and friction described by Amontons' law, the tensions just inside the ends of the arc are:

$$T_A = T_1\cos\theta_1 - \mu T_1 \sin\theta_1 \tag{14}$$

$$T_B = T_2\cos\theta_2 + \mu T_2 \sin\theta_2. \tag{15}$$

Now, on this arc AB it can be shown that

$$\frac{dT}{d\theta} + \mu T = N - \mu\frac{dN}{d\theta},$$

$$= 0 \tag{16}$$

as $N = 0$ on the arc. (The normal distributed load, w/unit length, can be shown to be given by $w = \frac{1}{\rho}\left(T + \frac{dN}{d\theta}\right) = \frac{T}{\rho}$ as $N = 0$.) So, from Eqn (16),

$$\frac{T_A}{T_B} = \exp\{\mu(\Phi - \theta_1 - \theta_2)\} \tag{17}$$

and combining Eqns (17) and (14) and (15),

$$\frac{T_1}{T_2} = \exp(\mu\Phi)\,\frac{\cos\theta_2}{\cos\theta_1}\,\frac{1+\mu\tan\theta_2}{1-\mu\tan\theta_1}\,\frac{\exp(-\mu\theta_2)}{\exp(\mu\theta_1)}. \tag{18}$$

If we assume that the friction at a point contact and the friction at an arc contact are both described by Amontons' law, but with different coefficients of friction μ_P and μ_A, then the equation for the ratio of tensions becomes:

$$\frac{T_1}{T_2} = \exp(\mu_A\Phi)\,\frac{\cos\theta_2}{\cos\theta_1}\,\frac{1+\mu_P\tan\theta_2}{1-\mu_P\tan\theta_1}\,\frac{\exp(-\mu_A\theta_2)}{\exp(\mu_A\theta_1)}. \tag{19}$$

Expanding the factor of $\exp(\mu_A\Phi)$ in Eqn (19) in powers of θ_1, θ_2, we obtain

$$\frac{T_1}{T_2} = \exp\mu_A\Phi\{1 - (\mu_A - \mu_P)(\theta_1 + \theta_2) + \tfrac{1}{2}(\mu_A - \mu_P)^2(\theta_1 - \theta_2)^2 - \tfrac{1}{2}(1+\mu_P^2)(\theta_2^2 - \theta_1^2) + \ldots\} \tag{20}$$

where, in Eqns (18), (19), (20), θ_1, θ_2 are given by Eqn (13).

Comments

With the coefficients of friction for concentrated and distributed loads the same, the correction to the standard capstan equation is of second order:

$$\frac{T_1}{T_2} = e^{\mu\Phi}\{1 - \tfrac{1}{2}(1+\mu_P^2)(\theta_2^2 - \theta_1^2)\}.$$

Here the decrease in the tension drop caused by the decrease in the angle of contact is being compensated (to the first order in θ_1, θ_2) by the friction forces associated with the

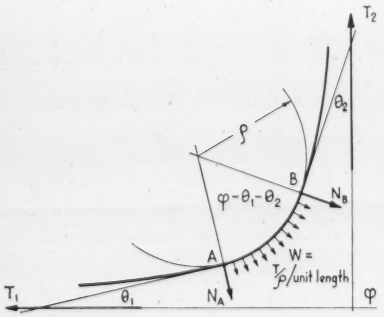


Fig. 3. Forces on a string in arc contact with a peg radius ρ.

concentrated reactions at the ends of the arc. The capstan equation is correct to 1% for $\mu < 1$ if the angles θ_2, θ_1 are less than 6°, or alternatively for the tensions T_1, $T_2 > 100\,T$, where $T = B/\rho^2$.

If the coefficient of friction varies with load, so that the coefficient of friction for a distributed load differs from that for a concentrated load, this compensation cannot take place, and there would be a rigidity correction to the capstan equation of the first order in θ_1, θ_2.

In this case the capstan equation could be in error by more than 10% for $T_2 < 4T$. For example Howell (1954) explains a dependence of T_1/T_2 on peg radius using a theory of coefficient of friction dependence on load. The nylon monofilament used had a rigidity modulus of the order of 5×10^2 dyn/cm^2 (for a radius 0·004 in. and assumed Young's Modulus $\sim 7 \times 10^{10}$ dyn/cm^2). The tensions T for the two pegs of radii 2·5 cm and 0·39 cm are then $T_{2\cdot5} = 0{\cdot}08$ g wt, $T_{0\cdot39} = 3{\cdot}3$ g wt.

The tension range of the smaller tension is $3 < T_2 \times 15$ g wt. So, for the smaller peg $T < T_2 < 5T$; and for the larger peg, $37T < T_2 < 200T$. Thus the results for the smaller peg are in the region where correction of the capstan equation for the rigidity of the fibres may be significant.

References

HOWELL, H. G., 1954, *J. Text. Inst.*, **45**, T575.

HOWELL, H. G., MIESKES, K. W., and TABOR, D., 1959, *Friction in Textiles* (London: Butterworth).

Notes and Comments

Russian translating programme

The Department of Scientific and Industrial Research are translating into English and publishing every month seventeen of the leading Russian technical journals. These provide up-to-date information on Soviet progress in steel manufacture, castings, welding, rubber and plastics, production engineering, machines and tooling, instrument construction, textiles, wood processing, chemistry, biology and mathematics.

The Department also sponsors the translation and publication of Russian technical books and each year places on sale translations of hundreds of technical articles. These translations are obtainable from the new National Lending Library for Science and Technology, Boston Spa, Yorkshire.

New British Standards

The following British Standards have recently been issued: British Standard 1991 : 1961, Recommendations for Letter Symbols, Signs and Abbreviations: Part 2, Chemical Engineering, Nuclear Science and Applied Chemistry; Part 3, Fluid Mechanics; Part 4, Structures, Materials and Soil Mechanics; Part 5, Applied Thermodynamics; British Standard 1828 : 1961, Reference Tables for Thermocouples (Copper *v.* Constantan); British Standard 3403 : 1961, Specification for Indicating Tachometers for General Industrial Use; British Standard 3385 : 1961, Specification for Direct-reading personal Dosemeters for X- and Gamma Radiation.

These may all be obtained from the British Standards Institution, British Standards House, 2 Park Street, London, W.1.

Lectures on applied physical chemistry for metallurgists

A course of ten lectures on applied physical chemistry for metallurgists will be given by Dr. J. Mackowiak in the Metallurgy Department, Battersea College of Technology, on Wednesdays from 6.30–9 p.m., commencing on 11th October 1961.

The course is designed to meet the needs of those employed on research and production who either completed formal studies some years ago or whose professional training did not include any previous study of the subject. Subjects to be covered come under the four main headings of Electron theory of elements and molecules, Thermodynamics, Kinetics and Application to practical problems.

Further information may be obtained from the Secretary (Metallurgy Courses), Battersea College of Technology, London, S.W.11.

Optical Industry Directory

The Optical Publishing Company, Lenox, Massachusetts, announces the publication of the seventh (1961) issue of the *Optical Industry Directory*. This gives a list of the thousand companies in America who manufacture instruments and components and supply services available from the optical industry. The lists are categorized under 350 item groupings. A revised lens appendix gives information concerning a thousand commercially available lenses for photography and optical instrumentation. The Directory may be obtained from The Optical Publishing Company, Box 1753, Lenox, Massachusetts. There are 200 pages. The price is $7.50.

Contact ferroelectric effect applied to plastic flow

by B. S. SATYANARAYANA, M.Sc., Faculty of Applied Physics, Institute of Armament Studies, Kirkee, Poona, India

MS. received 6th February 1961, in revised form 5th June 1961

Abstract

Investigations made on the ferroelectric effect of manganese dioxide show that a phenomenon, dependent on the efficiency of contact of the oxide surface with the electrodes, exists. The apparent changes in the ferroelectric hysteresis loops with pressure on the electrodes are traced experimentally. The apparent contact variation of the ferroelectric effect in manganese dioxide is applied to a comparison of the plastic flows in indentations in metals, the indentations being made by the Vickers hardness diamond pyramid.

1. Introduction

The ferroelectric effect in manganese dioxide has been investigated by Bhide and Damle (1960). It is found that the ferroelectricity vanishes at a temperature above 50° C, which is the Curie temperature of the material. The phenomenon is observed to be conspicuous at low temperatures and is supposed to vanish rapidly as the Curie temperature is reached. The investigations made here indicate that the apparent ferroelectric behaviour is dependent prominently on the nature of contact, and that the hysteresis loop area is very much reduced when the pressure of contact between the manganese dioxide and that of the electrodes pressing the manganese dioxide exceeds a certain limit. The apparent variation of ferroelectric hysteresis with the unidirectional pressure of electrodes causing different amounts of contact has been experimentally determined. The plastic flows in indentations of materials like steel, brass and other metals, during the process of indentation by a Vickers hardness diamond pyramid, are compared by tracing the apparent changes in the saturation polarization of ferroelectric effect with the efficiency of contact.

2. Experimental procedure

Chemically pure manganese dioxide powder, supplied by Thomas Tyers Co., was made into pellets by an application of pressure of 4000 lb/in^2. These pellets retained their compactness under the different experimental conditions. The pellets had a diameter of 2·12 cm and had thicknesses varying from 0·5 cm to about 1 cm. The pellet was placed between two flat carbon brushes, the contact surfaces being polished. Over the top brush, a suitable metal was inserted in order to indent it with the Vickers hardness diamond pyramid, as in Fig. 1. This arrangement enabled applications of standard loads determined by the Vickers hardness tester, so that the contact pressures between the pellet surfaces and the carbon brushes could be evaluated. The circuit employed for determining the ferroelectric polarization was that given by Sawyer and Tower (1930) and shown in Fig. 1. An alternating voltage of 20 v at 100 c/s (from a beat-frequency oscillator) was applied to the manganese dioxide tablet. The potential difference across AB was applied to the Y-plates of the cathode-ray oscillograph. A time-base voltage applied to the X-plates was obtained from the same beat-frequency oscillator. The apparent saturation value of the ferroelectric polarization was recorded for different loads conveyed by the Vickers diamond indenter.

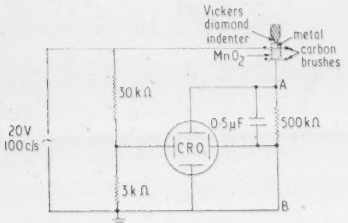


Fig. 1. Circuit for the study of the contact variation of ferroelectric effect in manganese dioxide at different pressures.

Experiments with aluminium and other metal electrodes give a slightly lesser output, possibly due to the metal oxides, which introduce extra resistance. The nature of the curves obtained with these electrodes conformed to the general trend shown in Fig. 4, but the deflection was comparatively small. Hence carbon brushes were chosen for the electrodes. Under high pressures, the manganese dioxide pellet has negligible resistance. The resistance across the carbon electrodes under the above conditions is the same as one would get when contact is made directly (between the carbon electrodes). This indicates that the carbon surface will maintain the best contact. The manganese dioxide pellet gives the maximum deflection in the cathode-ray oscillograph when its surface is in a fine powdery form, rather than in a glazy smooth condition. To ensure this every time, measurement is made after the surface has been cleaned with a cloth. Moisture or washing the surface with a mild acid makes the tablet almost conducting. Sparking at contact surfaces by high voltage application decreases the ferroelectric output.

3. Contact variation in ferroelectric loops

The manganese dioxide specimens used by Bhide and Damle (1960) were sintered, while the compacts employed here were not sintered for the study of the apparent ferroelectric changes. The typical ferroelectric hysteresis loops at low and high pressures are shown in Figs 2 and 3. The curve drawn in Fig. 4 shows the variation of ferroelectric saturation polarization with the unidirectional pressure applied on the contacts in the case of a number of tablets. The unidirectional pressure refers to the pressure acting on the manganese dioxide compact surface and is the ratio of

total load applied on the indenter to the area of the compact. The ferroelectric change would be just dependent on the pressure effective on the manganese dioxide compact surface. The saturation polarization is expressed in arbitrary units corresponding to the voltages impressed on the cathode-ray oscillograph plates for all the figures. It is found that the

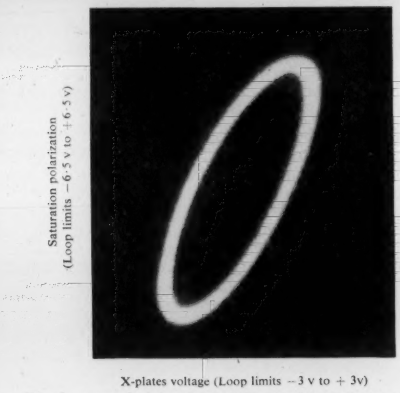


Fig. 2. Ferroelectric hysteresis loop at low-pressure contact.

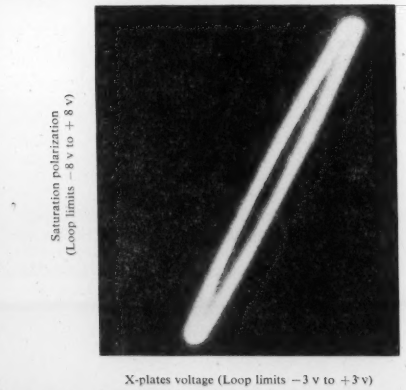


Fig. 3. Ferroelectric hysteresis loop at high-pressure contact.

ferroelectric output shows an exponential rise at small pressures, tending to a steady saturation value at high pressures. The contact ferroelectric hysteresis loop thereafter tends to decrease to almost negligible area with further increase of pressure, though the saturation polarization value does not decrease.

The apparent change in ferroelectric output may clearly be explained by the assumption that the fractional rate of increase with pressure of the number of molecules in contact is proportional to the fraction of molecules not yet in contact. The efficiency of contact at any stage is defined as the ratio of the saturation polarization P at that stage to the maximum saturation polarization P_{max} corresponding to the maximum amount of contact obtained by an application of very high pressure. An equation of the type

$$P = P_{max}\{(1 - e^{-cP})\}$$

where P is the pressure on the manganese dioxide compact surface and c is a constant, fits in with the experimental curve obtained in Fig. 4. The voltages corresponding to saturation polarizations at different stages are plotted in the figure. By analogy of time constants of electrical circuits, the reciprocal of constant c may be termed as the contact

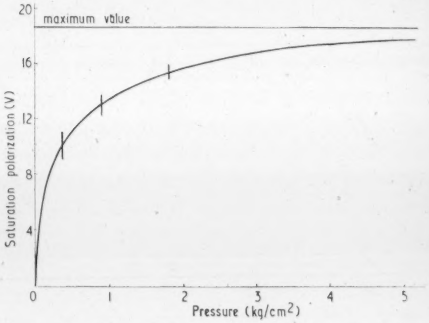


Fig. 4. Saturation polarization voltage in cathode-ray oscillograph plotted against pressure on manganese dioxide compact surface.

pressure constant (s or $1/c$). The value of s computed from the experimental curve for manganese dioxide is 0·5 kg/cm² at 25° C. It is found that the exponential rise of saturation polarization with pressure becomes steeper and reaches the maximum value for comparatively lower pressure at higher temperatures, of the order of 150° C. Obviously, the contact pressure constant decreases with temperature for manganese dioxide.

4. Plastic flow in materials

The contact variation of ferroelectric effect was utilized to determine the relative rates at which plastic flow took place in different materials, when the Vickers diamond pyramid indented the materials. As the indenter pierced materials like steel or brass, the ferroelectric saturation polarization of the manganese dioxide tablet placed beneath the metal increased gradually to a maximum steady value, the stage

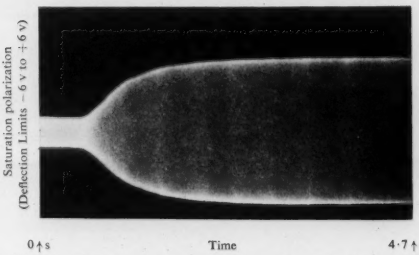


Fig. 5. Change of saturation polarization with time during indentation of steel by the Vickers diamond pyramid.

when the indentation is fully acting on the tablet surface. The saturation polarization was determined experimentally by removing the X-plates voltage in the cathode-ray oscillograph and tracing the variation of ferroelectric saturation polarization voltage only. The changes were recorded by a Cossor camera unit in which the film was moving at a regular speed of 0·25 in/s. The deflection increases gradually and the patterns obtained in the cathode-ray oscillograph for steel and brass are shown in Figs 5 and 6. From knowledge

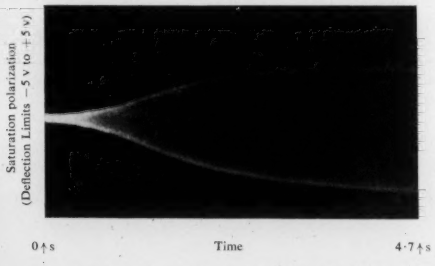


Fig. 6. Change of saturation polarization with time during indentation of brass by the Vickers diamond pyramid.

of the pressure acting on the manganese dioxide compact versus deflection curve shown in Fig. 4, and from the patterns in Figs 5 and 6, the variation of pressure on the manganese dioxide compact surface versus time curve could be computed, and is shown in Fig. 7. In this figure the limit of pressure is the ratio of the maximum load applied by the Vickers

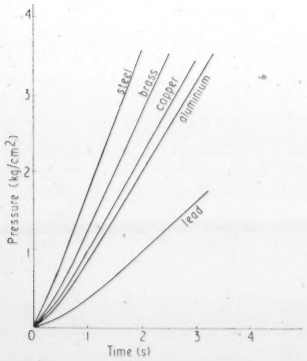


Fig. 7. Pressure on manganese dioxide compact surface during Vickers diamond pyramid indentation plotted against time curves. (Note: the maximum pressure reached is the maximum load applied by the Vickers machine divided by the area of the manganese dioxide compact surface.)

machine to the area of the tablet surface. The full load applied by the Vickers indenter is transmitted through the test metal on to the surface of the manganese dioxide compact and the pressure on the compact reaches a maximum after a few seconds. During this interval the plastic flow will be taking place in the indentation. The relative rates of plastic flow in indentations of materials, such as steel, copper, brass, aluminium and lead, can be qualitatively estimated by determining the slopes of the computed curves. The plastic flow is large for materials like lead. With any specific load, for a lead indentation compared to a steel one, the pressure on the manganese dioxide compact surface increases rather slowly to the maximum value. The rate of pressure increase on the compact surface (dp/dt) is found to increase with time. The initial part of the curve is non-linear, while the latter part is linear. The non-linear portion is small for steel and is large for a material like lead. The slope of the curve at any point is inversely proportional to the plastic flow rate. The experimental values of dp/dt are given in the Table.

Material	*Steady rate of increase of pressure on manganese dioxide compact surface* (dp/dt) (kg cm^{-2} s^{-1})
Steel	1·40
Brass	1·20
Copper	1·01
Aluminium	0·92
Lead	0·50

5. Conclusion

The apparent ferroelectric effect exhibited by manganese dioxide is found to be a contact dependent phenomenon, to a large extent. It increases rapidly with the efficiency of contact caused by the application of pressure. The apparent hysteresis loss tends to decrease at a very high pressure when contact is fully made. The efficiency of contact is assessed by measuring the apparent ferroelectric saturation polarization developed across a pellet of manganese dioxide and the variation of this saturation polarization with pressure is experimentally determined. An application of the contact variation of ferroelectric effect gives the relative plastic flow in indentations of materials made by the Vickers hardness indenters.

Acknowledgments

It is a pleasure to thank Air Commodore O. P. Mehra for permission to publish this material and Dr. B. Patnaik for his encouragement in the investigations. My thanks are also due to Mr. M. K. Subba Rao for valuable discussions and to Mr. T. P. S. Nair and to Mr. M. R. Katti for their assistance in experimental work.

References

Bhide, V. G., and Damle, R. V., 1960, *Physica*, **26**, 51.
Sawyer, C. B., and Tower, C. H., 1930, *Phys. Rev.*, **35**, 269.

Estimation of the total emission of a diode under normal operating conditions without removing the space-charge

by C. S. BULL, Ph.D., F.Inst.P., and R. K. FITCH, B.A., M.Sc., A.Inst.P., Department of Physics, Birmingham College of Advanced Technology, Gosta Green, Birmingham 4

MS. received 10*th March* 1961, *in revised form* 18*th April* 1961

Abstract

The total emission of the cathode of a planar diode over a range of cathode temperatures has been measured in a non-destructive way by making a detailed examination of the characteristics. A total emission of about 25 A/cm^2 *has been found for an EA* 50 *operating at normal cathode temperatures. This result has been tested by comparing theoretical values of the space-charge smoothing factor calculated from the characteristic curves in two distinct ways, only one of which involves the total emission directly.*

It was found that Richardson's equation for the total emission fails noticeably over the wide range of cathode temperatures investigated, and that this failure is not explicable on the basis of changes in the cathode work function.

Introduction

THE total emission i_s of a diode at normal operating cathode temperatures cannot be measured satisfactorily by d.c. techniques owing to the influence of the space-charge. Attempts have been made to overcome this difficulty by using pulse techniques (Sproull 1945, Coomes 1946, Wright 1949). At low cathode temperatures the effect of the space-charge can be neglected and the anode current saturates at very low anode potentials. However, it is useful to know the total emission for any high temperature, not only in the study of the emission itself, but also for the study of other properties such as the microphonic sensitivity and noise.

This paper reports a method of measuring the total emission without exceeding an anode voltage of more than 2 volts or an anode current of more than a few milliamps. The valve used was an almost planar commercial diode, the EA 50, and the cathode temperature was varied over a range from 680° K to 1130° K.

This method is based on theory given by Bull (1950) and is carried out as follows. The contact potential and its variation with cathode temperature is estimated from a set of characteristic curves of (log i, V) drawn for low cathode temperatures at which pronounced saturation of the emission occurs. The contact potentials so determined are then used in conjunction with the characteristic curves for higher and normal cathode temperatures to determine the total emission. The assumption needed in this method is that the trend of the contact potential at low cathode temperatures is maintained up to the normal operating temperature.

The validity of this assumption has been checked by calculating the space-charge smoothing factor of the diode in two ways based on the expression of the characteristic of the diode in terms of thirty partial differential coefficients connecting five interdependent variables (Bull 1950), namely the anode current density i, the total emission density i_s, the anode voltage V, the potential V_m of the barrier in the space-charge and the anode–cathode distance x.

One equation between these differential coefficients is

$$\left(\frac{\partial i}{\partial i_s}\right)_{Vx} = \frac{xg_x}{2i_s} + \frac{i}{i_s} \text{ where } g_x = \left(\frac{\partial i}{\partial x}\right)_{Vi_s}$$

from which, by putting $\xi = xg_x/2i$, we find

$$\left(\frac{\partial i}{\partial i_s}\right)_{Vx} = \frac{i}{i_s}(1+\xi). \tag{1}$$

The factor $1+\xi$ is found to occur frequently in the equations relating to the differential coefficients, and is also found (Bull 1950) to be a theoretical estimate of the space-charge smoothing factor, Γ^2.

If the values of i_s are known for two characteristic curves for which the difference of the cathode temperatures is small, the value of $1+\xi$ can be calculated, since the values of $(\partial i/\partial i_s)_{Vx}$, i and i_s can be found.

Also Bull (1950) has shown that

$$\left(\frac{\partial V}{\partial V_m}\right)_{ix} = \frac{1}{g_m}\left(\frac{ei}{kT} + \frac{xg_x}{2}\cdot\frac{e}{kT}\right)$$

where $g_m = (\partial i/\partial V)_{i_s x}$, the conductance of the diode. Using the expression for ξ given above, this becomes

$$\left(\frac{\partial V}{\partial V_m}\right)_{ix} = \frac{1}{g_m}(1+\xi)\frac{ei}{kT} \tag{2}$$

in which $(\partial V/\partial V_m)_{ix}$, g_m and T can be determined from the curves and e the electron charge and k Boltzmann's constant are known.

Using Eqns (1) and (2) we have two alternative methods of calculating $(1+\xi)$ for every point on the characteristic.

The cathode temperature for use in Eqn (2) is determined from the logarithmic relation between the anode current and the anode voltage in the retarding region of the characteristic,

$$i = i_s \exp_e(-eV/kT)$$

from which

$$\ln i = \ln i_s - \frac{e}{kT}V. \tag{3}$$

Provided that i_s is constant for any particular cathode temperature at all anode currents and voltages used, the slope of the graph of (log i, V) is equal to $e/2{\cdot}303\,kT$ and is therefore a measure of the cathode temperature. The linearity of the curves shown in Fig. 1, and very many other

observations of a similar kind, indicate that this method of finding T is at least as reliable as the methods commonly employed, and the temperatures so found do in fact correspond to the temperature at which electrons are emitted. Nergaard (1952), however, states that discrepancies are found

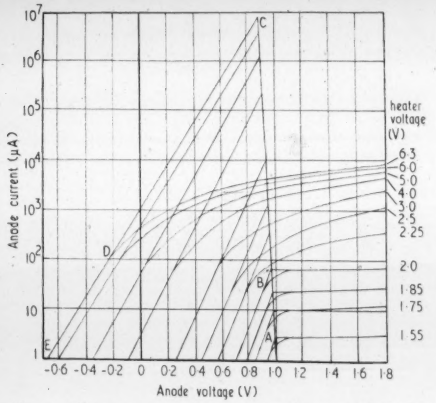


Fig. 1. Characteristic curves of log i plotted against V for EA50 at various heater voltages.

at the higher cathode temperatures between temperatures found this way and those obtained by using pyrometers.

Experimental method

In drawing the characteristic curves, filament voltages in the range 1·55 v to 6·3 v were used. The anode current was measured with a microammeter, and the anode voltage was varied over a range from −1·0 v to 2·0 v. It was found that it was possible to obtain reproducible and stable results only if the curve for the maximum cathode temperature used was drawn first, followed by those at lower temperatures. In this way the state of activation of the cathode and hence the emission was 'frozen' in as suggested by Metson *et al.* (1952). Considerable variations of the characteristics were found when recorded in the reverse order.

Experimental results and discussion

A typical set of the diode characteristics obtained is shown in Fig. 1. The temperature of the cathode was calculated from the slopes of the characteristics in the retarding region, from Fig. 1 using Eqn (3). The resulting variation of the cathode temperature with filament voltage is shown in Fig. 2. The shape of this curve is similar to one reported by Metson (1955) for a 6D15 diode, but in that case there was no indication of how the temperature was measured.

In Fig. 1 the 'knee' of the characteristic at low cathode temperatures shows a small but regular decrease of the cathode–anode contact potential as the temperature increases, as is shown by the slope of the line AB. This line AB, extrapolated to C intersects the extrapolation of the retarding region characteristics for the higher cathode temperatures, e.g. line EDC at C. For a cathode voltage of 6·3 v, for example, a total emission of about 8 A is observed, which corresponds to a total emission density of about 25 A/cm², since the coated area of the cathode is about 0·3 cm². This value is amongst the highest values quoted elsewhere. Metson (1955) reports a similar conversion method, but converts the emission in microamps at 700° K to the emission in A/cm² at 1020° K by using a factor which varies from valve to valve, but is of the order 1000. For the curves recorded here this factor would give a total emission density of possibly about 10 A/cm².

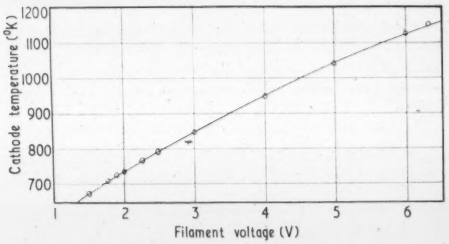


Fig. 2. Cathode temperature plotted against heater voltage.

The graph of log $(1 + \xi)$ derived from the characteristic curves is plotted as a function of the anode current in Fig. 3. The values obtained from using Eqn (1) are indicated by the circles and those from the use of Eqn (2) are indicated by triangles. It is seen that the two sets of points lie very close together, and this could be regarded as supporting the validity of the method of finding the total emission. Furthermore, this curve is in good agreement with a similar curve reported by Richards (1955).

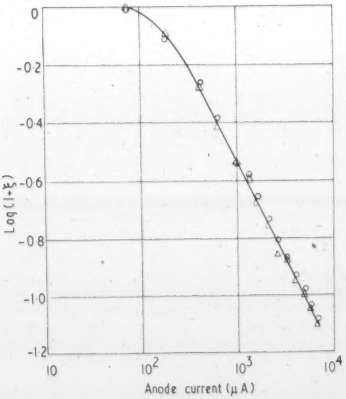


Fig. 3. Log $(1 + \xi)$ plotted against anode current.
O values using Eqn 1 Δ values using Eqn 2

From Fig. 1 it is also seen that the contact potential difference between the anode and the cathode decreases with increasing cathode temperature. If therefore we assume that the work function of the anode remains sensibly constant, this indicates that the cathode work function has a positive temperature coefficient of about 0·3 mv/deg K. This is a similar change to that reported by Hopkins and Vick

(1958) and by Young (1952), but is slightly smaller in magnitude. It also agrees with a result quoted by Bulyginski and Dobretsov (1956) who give a positive linear temperature coefficient of local work function for mixed oxide cathodes.

Using the values of total emission obtained from the curves, a graph of ($\log i_s/T^2$, $1/T$) was drawn, from which the applicability of Richardson's equation, $i_s = AT^2 \exp(-b/T)$, to the emission could be tested. In the equation A and b are usually considered to be constants independent of temperature. The resulting graph is shown in Fig. 4. This is not a straight line, as would have been expected if the quantities A and b were in fact constant.

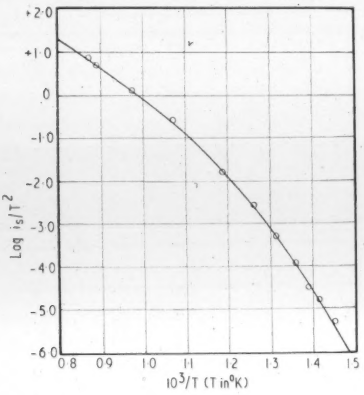


Fig. 4. Richardson plot.

If it be assumed that A is constant, then the curve indicates a very pronounced fall in the value of b, i.e. the work function, as the temperature increases. On the other hand, the line AB in Fig. 1 indicates that the contact potential difference between the cathode and the anode does not vary very much. It is unlikely since the cathode and the anode are at different temperatures and consist of different materials that this could be the result of almost exactly the same variation taking place at the anode as at the cathode. It can be conjectured from Fig. 1 that the work functions do not vary very much with temperature. Consequently we might conclude from Fig. 4 that the quantity A, usually regarded as a constant, does in fact vary rapidly with temperature. This variation in the value of A could be attributed to evaporation or diffusion away from the surface of active centres responsible for the emission. Childs (1911) has already suggested that A is a function of temperature. Experiments using microsecond pulses, however, have produced linear Richardson plots over ranges of high cathode temperatures (Sproull 1945, Coomes 1946, Wright 1949).

It is possible that the non-linearity demonstrated in Fig. 4 is more clearly seen on account of the fact that the temperature range reported here is larger than that usually used. Close inspection of similar curves given by Higginson (1958), Hopkins and Vick (1958) and Davies and Hopkins (1959) show that in most instances a similar curvature is observed, but on account of the smaller temperature range it is not so easily seen. Graphs such as Fig. 4 are not very sensitive to the index of T used in the Richardson equation, so that it is not possible to explain the curvature by using indices other than 2, such as 5/4 or 1/2 as is sometimes done.

Conclusions

This work has demonstrated that the total emission of a diode can be determined using normal d.c. techniques even when the effect of the space-charge cannot be neglected, and without even momentarily overloading the valve. It is hoped to apply this method to other experimental diodes which are being made in this laboratory.

The results bring into question the possibility that the constants in Richardson's equation, which are assumed to be constant for the purpose of calculating the values of work functions from the emission, are not in fact even approximately constant for an oxide cathode. The work function of such cathodes cannot therefore be measured in this way. Hopkins (1960, Hopkins and Vick 1960) has applied the Kelvin vibrating electrode method with considerable success. However, even in this way absolute values of the work function may not be obtained on account of contamination of the reference electrode.

Acknowledgments

The authors wish to thank Mr. G. V. Cochrane for his assistance in taking some of the measurements and Mr. E. Robinson for help in the preparation of the paper.

References

Bull, C. S., 1950, *Proc. Instn Elect. Engrs*, **97,** 159.

Bulÿginskiĭ, D. G., and Dobretsov, L. N., 1956, *Zh. Tekh. Fiz.*, **26,** 1141.

Childs, C. D., 1911, *Phys. Rev.*, **32,** 492.

Coomes, E. A., 1946, *J. Appl. Phys.*, **17,** 647.

Davies, D. E., and Hopkins, B. J., 1959, *Brit. J. Appl. Phys.*, **10,** 498.

Higginson, G. S., 1958, *Brit. J. Appl. Phys.*, **9,** 106.

Hopkins, B. J., 1960, *Brit. J. Appl. Phys.*, **11,** 124.

Hopkins, B. J., and Vick, F. A., 1958, *Brit. J. Appl. Phys.*, **9,** 257.

—— 1960, *Brit. J. Appl. Phys.*, **11,** 223.

Metson, G. H., 1955, *Proc. Instn Elect. Engrs*, **102,** 657.

Metson, G. H., Wagener, S., Holmes, M. F., and Child, M. R., 1952, *Proc. Instn Elect. Engrs*, **99,** 69.

Nergaard, L. S., 1952, *R.C.A. Rev.*, **13,** 464.

Richards, C. B., 1955, M.Sc. Thesis (Wales).

Sproull, R. L., 1945, *Phys. Rev.*, **67,** 166.

Wright, D. A., 1949, *Proc. Phys. Soc.*, B, **62,** 398.

Young, J. R., 1952, *J. Appl. Phys.*, **23,** 1129.

Hypothesis on the ionic mechanism of the growth of whiskers obtained by reduction of the metal halides

by J. MAZUR, D.Phil., F.Inst.P., and J. RAFAŁOWICZ, M.Sc., Low Temperature Laboratory, Institute of Physics, Polish Academy of Sciences, and Physics Department of the University, Wroclaw, Poland

MS. received 14*th February* 1961

Abstract

On the basis of experimental results the authors put forward the hypothesis of the existence of an ionic mechanism of whisker growth side by side with a non-ionic (atomic or molecular) mechanism.

The ionic concentration calculated from the current density is of the order of 10^{12} *ions/cm*^{3} *at* 650° C *and therefore the possibility of an ionic mechanism of whisker growth cannot be excluded: this conclusion is reinforced by the observed velocity of growth of copper whiskers, and the influence exerted on this growth by an electric field. These preliminary measurements and qualitative observations make the simultaneous coexistence of the two growth mechanisms an acceptable hypothesis.*

THERE are at present two different views concerning the growth of copper whiskers at high temperature in a hydrogen atmosphere. The earlier conceptions accepted the uniform mechanism of growth (Sears 1953, 1959, Brenner 1956). According to this theory copper whiskers grow by condensation of the saturated copper vapour on the tip of the whisker; this theory thus explains the observed growth of the monocrystals from the top.

The equation of material transport by diffusion from the gaseous phase may be written

$$P = r\rho i 2\lambda_s (2\pi RT/M)^{1/2} \tag{1}$$

where P is the minimal pressure of copper vapour, r the whisker radius, ρ the whisker density, i the observed axial growth rate, λ_s the mean free path of surface diffusion of self-adsorbed copper atoms, R the gas constant, T the temperature on the Kelvin scale, and M the atomic or molecular weight.

Using this formula Brenner (1959) estimated the required copper vapour pressure necessary for the velocity of the whisker growth to agree with his experimental results (about 70 μ/s under optimum conditions of experiment).

From this calculation it follows that the required vapour pressure of copper should be of the order of 2×10^{-3} mm Hg (Morelock and Sears 1959).

In the experiments carried out by Morelock and Sears (1959) such a concentration of copper vapour was not found and the authors were compelled to postulate another quite different mechanism of growth. This second idea about the mechanism of whisker growth was that the vapour of the CuI was the carrier of the metal necessary for forming the copper whiskers. The molecules of the CuI undergo catalytic decomposition into Cu and I at the dislocation step on the top of the whisker; iodine reacts with hydrogen and the copper forms the material necessary for the growing monocrystal.

The experiments carried out by the present authors have revealed a considerable ion concentration of Cu^+ and I^-, and therefore suggests the hypothesis of the existence, among other possibilities, of an ionic mechanism of copper whisker growth. Measurements show that the density of the ionic current of the CuI vapour at the temperature 650° C in air is equal to $1{\cdot}4 \times 10^{-5}$ A cm^{-2} for an electric field 11·25 V cm^{-1}. These measurements were carried out in a glass tube with tungsten electrodes, and the applied electric fields were too weak to influence the ionic production, and sufficient only to affect the directional movement of the ions in the gaseous medium. The details concerning the apparatus and the experiments will be given in a subsequent paper.

The density of the ionic current found enables the ionic concentration in the space between the electrodes to be calculated as follows:

The current density i may be written

$$i = \sigma E, \tag{2}$$

Since

$$\sigma = n_i e u \tag{3}$$

$$i = n_i e u E \tag{4}$$

where σ is the electric conductivity, E the electric field, n_i the concentration of the current carriers, e the carrier charge, and u the ionic mobility.

The mobility u may be represented on the basis of the microscopic theory of ions mobility in the gas by the formula

$$u = 0{\cdot}64 e\bar{\lambda}/m\bar{v} \tag{5}$$

where e is the ionic charge, $\bar{\lambda}$ the mean free path of the ions, m the mass of an ion, and $\bar{v}$ the mean velocity.

It is well known that

$$\bar{\lambda} = 1/2^{1/2}\pi n_c d^2 \tag{6}$$

where n_c is the concentration of the gaseous medium and d the diameter of the molecules in the gaseous medium, and

$$\bar{v} = (8kT/\pi m^*)^{1/2} \tag{7}$$

where k is Boltzmann's constant, T the temperature on the Kelvin scale, and m^* the atomic or molecular mass.

Substituting the values (6) and (7) in (5) we get

$$u = \frac{0{\cdot}16e}{n_c d^2 (\pi m k T)^{1/2}} \text{ if } m = m^*. \tag{8}$$

Substituting (8) in (4) gives

$$i = \frac{0{\cdot}16 n_i e^2 E}{n_c d^2 (\pi m k T)^{1/2}} \tag{9}$$

where m is ionic mass, whence

$$n_i = \frac{d^2(\pi mk)^{1/2} i n_c T^{1/2}}{0{\cdot}16 e^2 E}. \qquad (10)$$

The formulae (2), (3) and (4) may be applied only in the case of free charges, while formula (5) is valid only for the free charges and weak electric fields (Kapcow 1950).

The assumption of weak electric fields enables formulae (6) and (7) to be used to get $\bar{\lambda}$ and $\bar{v}$; these formulae are valid generally for atoms and molecules.

The meaning of the final formulae (8) and (9) is detailed below. The concentration of copper ions estimated on the basis of the formula (10) equals 10^{12} ions cm^{-3}, and this corresponds to an ionic pressure of copper of $8{\cdot}4 \times 10^{-5}$ mm Hg.

This pressure is lower (by about one order) than the value 2×10^{-3} mm Hg required for a growth velocity 70 μ/s due exclusively to an ionic mechanism of growth. However, the optimal value of the growth velocity 70 μ/s, given by Brenner, is about one order too high by comparison with our results.

Here the mean length of the whiskers formed during the reaction lasting one hour was 10 mm, so that the velocity of the growth was about 3 μ/s.

Brenner (1959) during his study of the velocity of the copper whisker growth as a function of temperature, has applied very short reaction time and considered the longest whiskers, grown during this short period; thus he estimated the maximum of the copper whisker growth velocity under optimal conditions at the temperature 650° C as 70 μ/s.

We have estimated the mean copper whisker growth velocity during the whole period of the chemical reaction. Assuming that the mean growth time is equal to the reaction time (under our conditions about 1 hour) and the mean whisker length is about 10 mm, we get, as a mean growth velocity at the temperature 650° C, 3 μ/s.

For the time being it is not known whether the velocity is constant during the whole period of the whisker growth and whether the mean growth time is equal to the reaction time; it is therefore rather difficult to decide which of the two different approaches is more reasonable. The vapour pressure required for such a growth velocity is, according to Eqn (1), 10^{-4} mm Hg, and this is in agreement with our value $8{\cdot}4 \times 10^{-5}$ mm Hg. The additional directional action of the local electric field, generated by the ionic charges, also seems to indicate the possibility of an ionic mechanism of copper whisker growth.

The growth velocity may vary with slight changes of the vapour concentration, and the value of concentration obtained for the copper ions, namely, $8{\cdot}4 \times 10^{-5}$ mm Hg, seems to be quite reasonable.

The authors point out that the experiment carried out by Morelock and Sears (1959) on the estimation of the copper vapour concentration excluded the possibility of the presence of saturated copper ion vapour, because the Cu^+ and I^- ions in the vicinity of the surface of the cooled condensation tube associated into molecules of CuI and a preponderant quantity of those molecules was found during the experiments.

It would be difficult by using the experimental method, applied by Morelock and Sears (1959) to prove the presence of the ions in the CuI vapour. These authors postulated only a non-ionic mechanism of whisker growth, and did not consider the thermal dissociation of the CuI vapour as a result of the catalytic action of the hydrogen; it follows from our experiments that this fact plays an important part. The analysis of the deposit condensed on the molybdenum tube was not sufficiently accurate to enable final conclusions concerning the composition of the medium in the reaction and whisker growth zone to be drawn. Morelock and Sears (1959) stressed the complete lack of copper deposit on the walls of the silica tube in the reaction zone; this is also at variance with our findings and results. We found that the quantity of the copper deposit on the silica tube walls increased markedly when whisker growth occurred in the electric field.

In the paper of Morelock and Sears it is stated that the composition of the condensate on the sampling tube was independent of the furnace temperature and depended only upon the temperature of the sampling tube. In the view of the present writers this statement shows that the method of analysis of the deposit in the reaction and whisker growth zone was inadequate.

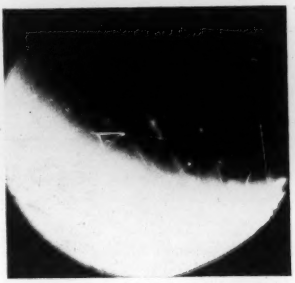


Fig. 1. Whisker growth without electric field.

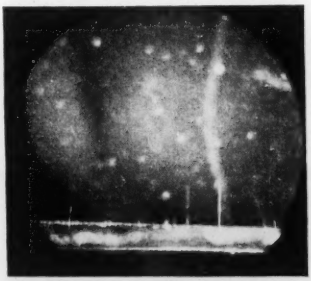


Fig. 2. The whiskers following the electric field.

The large concentration of the supersaturated vapour of copper ions throws new light on the possibility of one additional mechanism of metal whisker growth, i.e. the reduction of the halides.

The influence of the electric field on whisker growth during the reduction of the halides found during our investigation, supplies the direct proof of the above hypothesis. The influence of the electric field on the whisker growth, which caused a directional distribution of whiskers along the force lines, has been observed beyond doubt.

Figures 1 and 2 show the influence of the electric field on

copper whisker growth, Fig. 1 without electric field, Fig. 2 with electric field.

More details and the main experimental results will be published in a later paper on the influence of the electric field on copper whisker growth (Hoffmann *et al.* 1961). Side by side with the distinctly directed whiskers there are also random whiskers, and this may be a direct proof of the coexistence of the ionic and non-ionic atomic or molecular whisker growth mechanisms. The single whiskers on the upper electrodes formed as a result of material transport from the boat filled with CuI, and the applied electric field was 3 kv. It should be stressed that whiskers on the upper electrodes were obtained both when the applied potential was positive and when it was negative. This can be easily explained when one takes into consideration the existence of both Cu^+ and CuI^- ions, which undergo on the tip of the whisker a neutralization and catalytic decomposition into Cu and I; this process has been proved by Coleman and Sears (1957). The existence of the negative ions CuI^- during the dark electric discharge under atmospheric pressure is possible, according to data in the literature (Kapcow 1950).

The results of our experiments thus seem to prove also an ionic mechanism of the whisker growth; the full report will be published in the near future.

References

Brenner, S. S., 1956, *Acta Met.*, **4**, 62.

—— 1959, *Acta Met.*, **7**, 519

Coleman, R. V., and Sears, G. W., 1957, *Acta Met.*, **5**, 131.

Hofmann, T., Mazur, J., Niklibore, J., and Rafałowicz, J., 1961, *Brit. J. Appl. Phys.*, **38**, 342.

Kapcow, N. A., 1950, *Electric Phenomena in Gas and Vacuum* (in Russian) (Moscow: Government Edition of Technical and Theoretical Literature), p. 242.

Morelock, C. R., and Sears, G. W., 1959, *J. Chem. Phys.*, **4**, 926.

Sears, G. W., 1953, *Acta Met.*, **1**, 457.

—— 1959, *Acta Met.*, **3**, 361.

Notes and comments

Welding Engineering 1962

Welding Engineering 1962 is the first large-scale exhibition to be held by the Institute of Welding. It will take place at Buxton from 30th April to 4th May 1962, in connection with the Institute's Spring Meeting the theme of which will be Welding for Power Generation. The exhibition will illustrate the rapid progress of welding technology in this country, especially in meeting the challenge of modern power generation plant construction. There will be special sections devoted to Research, Education and Publications.

Further information may be obtained from the Institute of Welding, 54 Princes Gate, Exhibition Road, London, S.W.7.

IFAC Congress 1963

The second congress of the International Federation of Automatic Control (IFAC) will be held in Basle, Switzerland, in September 1963, at the invitation of the Swiss Association for Automatic Control. A committee of the British Conference on Automation and Computation will co-ordinate the submission of papers from the United Kingdom to the congress.

Most of the papers will deal with the theory or application of automatic control; a number will deal with the components of control devices. Subjects to be covered are as follows: Theory: discrete systems, stochastic systems, optimal systems, learning systems, systems reliability; Applications: process dynamics, computer studies of applications (on or off line), optimizing or adaptive control applications; Components: new and effective devices, measurement of the reliability of components.

Papers may be submitted in English, Russian, French or German. The final version of any paper by a United Kingdom author should be in the hands of the BCAC committee not later than 1st June 1962.

Offers of papers from the United Kingdom should be made to the Honorary Secretary, BCAC, c/o The Institution of Electrical Engineers, Savoy Place, W.C.2. General enquiries concerning the congress arrangements should be made to the Secretary of IFAC, Dr.-Ing. G. Ruppel, Prinz-Georg-Strasse, 79, Dusseldorf, Germany.

The Physical Chemistry of Aerosols

We have received a copy of The Physical Chemistry of Aerosols, Discussions of the Faraday Society, No. 30, 1960. This discussion was held in the Chemistry Lecture Theatre, Queen's Building, Bristol University, on 13th, 14th and 15th September 1960. The papers are divided into three sections: Nucleation, homogeneous and heterogeneous; Growth of particles; Physical and chemical properties.

It is published by The Faraday Society, 6 Gray's Inn Square, London, W.C.1. The price is 50*s*.

Experimental estimation of the degree of the ionic thermal dissociation of the CuI vapour in the air atmosphere between 350 °C and 550 °C

by J. MAZUR, D.Phil., F.Inst.P., and J. RAFAŁOWICZ, M.Sc., Low Temperature Laboratory, Institute of Physics, Polish Academy of Sciences, and Physics Department of the University, Wrocław, Poland

MS. received 26*th April* 1961

Abstract

On the basis of the measurements of the ionic current density in the CuI vapour the ionic pressure has been evaluated as a function of temperature and afterwards compared with data given in the literature, concerning the total pressure of the CuI vapour.

Thus it was possible to calculate the thermal ionic dissociation of the CuI vapour as a function of temperature and to plot the result.

As a first step in our investigation on the mechanism of the metal whisker growth by using Brenner's (1956, 1959) method the following experiment has been carried out.

A quantity of CuI was put in a glass tube made from very hard glass, Termisil, fitted with tungsten electrodes; the tube was inserted in a tube furnace and connected to an electric circuit containing a galvanometer and battery (Fig. 1).

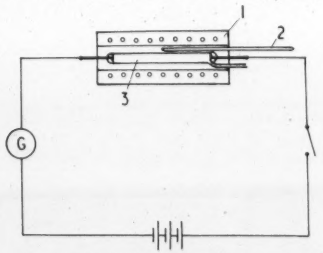


Fig. 1. Circuit diagram of the apparatus.
1. Electrical tube furnace.
2. Thermometer.
3. Tube with electrode.

One end of the tube was left open in order to ensure the same pressure inside as during the whisker growth using Brenner's method. Thus it was possible not only to confirm the existence of the ions in the tube, but also to measure the ionic current as a function of temperature (Fig. 2).

The measured value of the ionic current enables the ion concentration n_i in the space between the electrodes (Mazur, Rafałowicz, to be published) to be calculated, using the formula

$$n_i = \frac{d^2 i n_c (\pi m k T)^{1/2}}{0{\cdot}16 e^2 E} \tag{1}$$

where d is the diameter of the molecule in the gaseous medium, m the ion mass, k Boltzmann's constant, e electron charge, E the electric field between the electrodes, n_c the concentration of all the molecules of the gaseous medium, T temperature (° K) and i density of the ionic current. The pressure

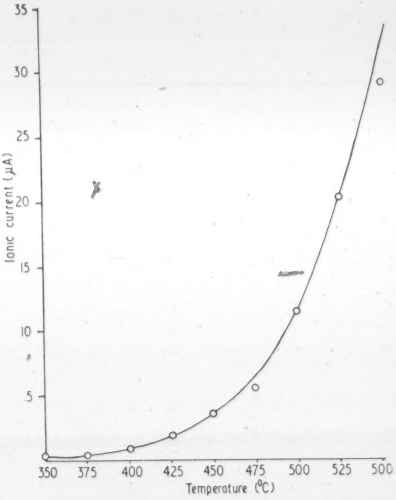


Fig. 2. The ionic current as a function of temperature.

inside the tube is the same as in the ambient air p_{atm} because the tube is open. Thus

$$p_{atm} = n_c k T. \tag{2}$$

Substituting for n_c in (1) we get

$$n_i = \frac{d^2 p_{atm} (\pi m/k)^{1/2} i}{0{\cdot}16 e^2 E T^{1/2}}. \tag{3}$$

The partial pressure of the ionic vapour may be written as

$$p_i = n_i k T \tag{4}$$

whence, from (3),

$$p_i = \frac{d^2 p_{atm} (\pi m k T)^{1/2} i}{0{\cdot}16 e^2 E}. \tag{5}$$

Thus the ionic vapour pressure p_i may be evaluated as a function of temperature, if the density of the ionic current is known.

The relation between p and T for CuI, given approximately by the formula

$$\log p = a - \frac{b}{T} \qquad (6)$$

is represented on the graph by a straight line (Fig. 3, A). If we plot on the same graph the values log p_i, obtained from our measurements, we also get a straight line, B.

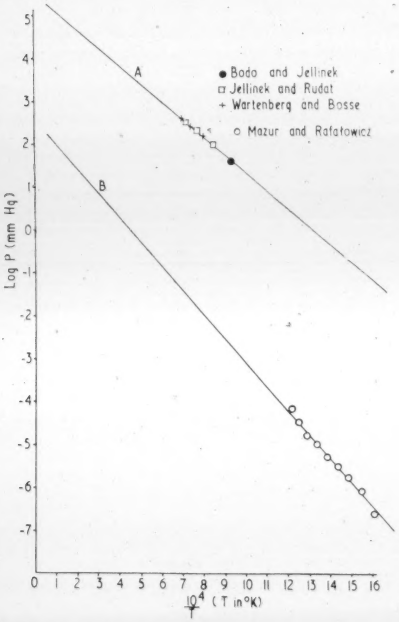


Fig. 3. Log of the CuI vapour pressure as a function of temperature.

Thus it is clear that the increase of the ionic vapour pressures Cu^+ and I^- is governed by the same rule as the increase of the total vapour pressure CuI.

Now from Eqns (2) and (4) $p_i/p_{atm} = n_i/n_c = \alpha_i$ say, where α_i will denote the degree of ionic thermal dissociation. Then from Fig. 3 which gives graphs of log p_i and log p_{atm} we can at once determine log α_i. For example for $T = 10^4$ °K log $p_i \simeq 2$, log $p_{atm} \simeq 5$ so that log $\alpha_i \simeq -3$ and $\alpha_i \simeq 10^{-3}$.

If we plot α_i against $1/T$ the curve so obtained is also a straight line (Fig. 4). Fig. 4 contains the results of our measurements.

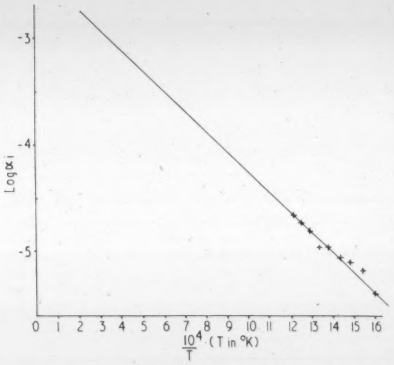


Fig. 4. Log of the degree of ionic thermal dissociation as a function of temperature.

The CuI vapour subsequently dissociates thermally into the atoms Cu and I at a temperature of 650° C, but not into the Cu^+ and I^- ions, because the atomic dissociation is, on the basis of the energy consideration, more probable than the ionic dissociation; this has been confirmed as far as NaCl is concerned (Kondratjew 1959). The formation of Cu^+ and I^- as a result of thermal dissociation or evaporation seems rather improbable.

The results of our measurements enable us to estimate the degree of the ionic thermal dissociation of the CuI vapour as a function of temperature. CuI^- ions cannot be excluded from our experiments.

Thus the Cu^+ and I^- ions existing in the gaseous medium are for the most part generated in a way different from the thermal vapour dissociation of CuI molecules.

We cannot exclude the possibility that some other process is operative, for instance catalytic or chemical influence of the gaseous medium.

The problem still remains unsolved and further investigations are in progress.

References

BRENNER, S. S., 1956, *Acta Met.*, **4**, 62.

—— 1959, *Acta Met.*, **7**, 519.

GREINER, B., and JELLINEK, K., 1933, *Z. Phys. Chem.*, A, **165**, 117.

JELLINEK, K., and RUDAT, A., 1929, *Z. Phys. Chem.*, **143**, 58.

KONDRATJEW, B. H., 1959, *Atomic and Molecular Structure* (in Russian) (Moscow: Government Edition of Technical and Theoretical Literature), p. 292.

WARTENBERG, H., and BOSSE, O., 1922, *Z. Elektrochem.*, **28**, 384.

Optical absorption of some oxides in the Schumann ultra-violet region

by W. P. DOYLE, Ph.D., D.I.C., University College, Dublin, and F. J. P. CLARKE, Ph.D., A.Inst.P., Atomic Energy Research Establishment, Harwell, Berks.

MS. received 11*th April* 1961, *in revised form* 26*th May* 1961

Abstract

The optical absorption of thin films of stannous oxide, stannic oxide, lead monoxide, the trioxides of arsenic, antimony and bismuth, beryllium oxide, zirconium dioxide and tantalum pentoxide has been measured in the energy range 4 *to* 10·5 *ev.*

Introduction

Except for magnesium oxide (Reiling and Hensley 1958), little is known about the optical absorption of oxides in the Schumann ultraviolet range. Prior to detailed study of some oxides in this region, a survey has been made of optical absorption in a number of oxides, most of which have previously been investigated above 2000 Å.

Experimental details

Specimens

For the determination of the fundamental optical absorption in solids, films of the materials a few hundred angstroms thick are necessary. The requisite thin films were mostly prepared by vacuum evaporation using as supports freshly cleaved lithium fluoride plates 1–2 mm thick. Lithium fluoride was used as substrate because it is transparent out to 11·5 ev.

The films were prepared as follows: SnO, by heating a vacuum evaporated film of tin in air at 400° C for 15 minutes (Doyle and Lonergan, unpublished work); SnO_2, by reactive sputtering of tin in oxygen at 0·06 mm pressure (Doyle and Lonergan, unpublished work); PbO, by heating a vacuum evaporated film of lead in air at 650° C for 15 minutes (Doyle and Lonergan, unpublished work); As_2O_3, by direct evaporation (Doyle 1958); Sb_2O_3, by direct evaporation (Doyle 1958); Bi_2O_3, by heat oxidation of a vacuum evaporated bismuth film (Doyle 1958); BeO, by evaporation of beryllium from a tantalum strip heater and heating the beryllium film so obtained in air at 600° C for 30 minutes; the resulting beryllium oxide film was uniform, white and slightly turbid; ZrO_2, by vacuum evaporation of the material from a tungsten heater (Holland 1956); a clear colourless film was obtained; Ta_2O_5, by vacuum evaporation of the material from a tungsten heater (Holland 1956); a clear colourless film was obtained.

All the substances studied are chemically stable in air at room temperature and so it was not necessary to vacuum deposit a protective coating of lithium fluoride or other material on to the films.

Measurement of absorption

The absorption spectra of the films from 4 to about 10·5 ev were determined using the Schumann region grating spectrophotometer previously described (Clarke and Garton 1959). The light source used was the positive column of a hydrogen discharge tube run at 400 mA. The discharge tube was not sealed off from the spectrometer by a lithium fluoride plate (Clarke and Garton 1959) so that the spectrometer was in direct communication with the discharge tube and consequently filled with hydrogen at 1–2 mm pressure. The radiation detector employed was a salicylate-sensitized E.M.I. Type 6256 photomultiplier. Optical density (log I_0/I_t) was determined by measuring the microammeter deflection with the absorbing specimen out of the beam I_0 and the deflection with the specimen and lithium fluoride mount in the beam I_t. The deflection due to the photomultiplier dark current was balanced out before each I_0 and each I_t reading.

It will be apparent that the method of measuring the absorption of the solid films did not allow for the absorption of the lithium fluoride substrate. The absorption of a typical blank lithium fluoride plate is shown in Fig. 1. It

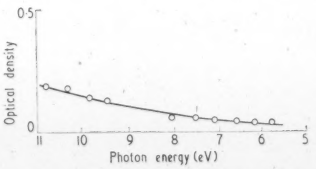


Fig. 1. Absorption spectrum of lithium fluoride plate.

can be seen that the only effect of the absorption of the substrate will be to raise slightly the absorption curve of the film plus substrate from 9·5 to 11 ev.

In order to measure absorption at wavelengths greater than 2200 Å it was necessary to consider the possibility of effects due to second order spectra. The band spectrum of hydrogen extends from about 1700 to about 1100 Å; thus it may appear in second order from 3400–2200 Å. A scan of the region 3400–2200 Å showed that the most intense peaks of the band spectrum appeared with appreciable intensity in second order superimposed on the continuum and would therefore invalidate absorption measurements in this region. Insertion of a quartz plate in the incident beam, by absorbing most of the band spectrum, completely eliminated the second order peaks from 2200 to 3100 Å and all measurements at wavelengths greater than 2160 Å were made with the quartz plate in the incident beam. Presumably absorption measurements could be extended to wavelengths greater than 3100 Å by using, for example, a glass instead of a quartz filter.

In measuring I_0 (absorbing sample out of the beam) the ratio (u/s) of useful to stray light signals varied between

1700 : 1 and 130 : 1 depending on the wavelength: so that in general it was not necessary to correct for stray light in measuring I_0. However, in measuring I_t (absorbing sample in the beam) in a region of strong absorption, the u/s ratio will be much smaller and it is necessary to correct for stray light. A rough analysis of the spectral distribution of the stray light was made by using glass and quartz filters. It was shown that about 30% of the stray light lay at wavelengths shorter than 2700 Å, the rest lying at longer wavelengths. The absorption thresholds of the solids studied are such that most of the stray light below 2700 Å will be absorbed by the specimen and thus correction should be made only for the stray light of wavelength greater than 2700 Å. The latter was determined by setting the wavelength-micrometer to a wavelength outside the emission region of the source and using a glass filter, the absorption threshold of which was approximately 2700 Å, in the exit beam. Thus, assuming that the stray light level is independent of the wavelength setting of the instrument, the I_t values could be corrected for stray light. The fact that in regions of high absorption a significant part of the I_t signal may be due to stray light places a limit on the I_0/I_t ratios which can be accurately measured; in certain wavelength regions I_0/I_t ratios greater than 50 cannot be accurately measured. This difficulty could however always be overcome by the use of thinner films thus increasing the u/s ratio in measuring I_t.

In measuring the absorption of a film no corrections were made for reflection and scattering losses. For the determination of the shape and position of the absorption bands this is unimportant, the only disadvantage being that, in the long-wave part of the spectrum, absorption may be simulated due to reflection and scattering losses (Doyle 1958). No correction was made for possible fluorescence of the sample.

All measurements were taken with a band pass of 3 Å and absorption was measured at intervals of approximately 0·1 ev throughout the spectral region studied.

The performance of the spectrophotometer in measuring absorption was tested both above 2000 Å and in the Schumann region. Above 2000 Å the absorption spectra of stannous oxide, lead monoxide and the trioxides of arsenic, antimony and bismuth agreed with previous work (Doyle and Lonergan, unpublished work, Doyle 1958) on these materials using a Beckmann Model DU spectrophotometer. In the region 2000–1200 Å the absorption spectrum of a magnesium oxide film showed a sharp threshold at 6·9 ev and a peak at 7·7 ev in substantial agreement with the work of Reiling and Hensley (1958).

Results

The absorption spectra from 4 to about 10·5 ev of the substances studied are shown in Figs 2, 3 and 4. Optical

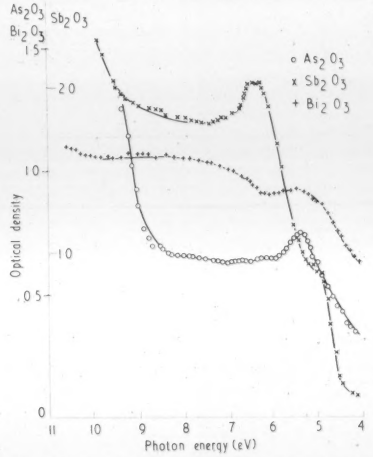


Fig. 3. Absorption spectra of thin films of arsenic, antimony and bismuth trioxides.

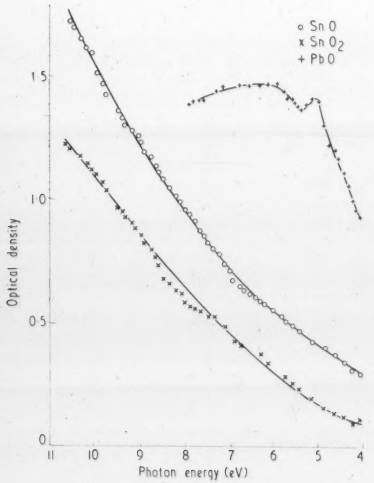


Fig. 2. Absorption spectra of thin films of stannous oxide, stannic oxide and lead monoxide.

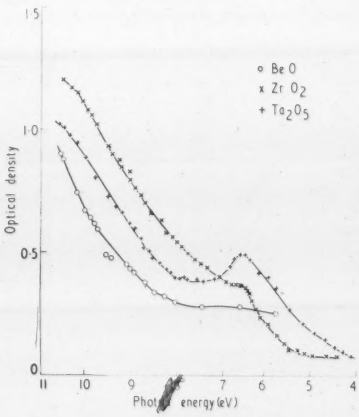


Fig. 4. Absorption spectra of thin films of beryllium oxide, zirconium dioxide and tantalum pentoxide.

density is shown as a function of photon energy. The thickness of the films has not been determined and so it is not possible to give accurate values of absorption coefficients. However, previous measurements (Doyle and Lonergan, unpublished work, Doyle 1958) above 2000 Å of absorption coefficients for some of the materials combined with the amount of substance evaporated permit the conclusion that the thickness of all the films is such that in the regions of maximum absorption the absorption coefficient is about 10^5 cm^{-1}.

Acknowledgments

We should like to thank Dr. W. R. S. Garton of Imperial College, London, for much helpful advice on the spectrophotometric measurements, Mr. G. A. Lonergan of Trinity College, Dublin, for the preparation of most of the films and Mr. J. Williams, Mr. N. Hill and Mr. R. G. Bellamy for advice on the evaporation of beryllium.

One of us (W. P. D.) expresses his sincere appreciation of the facilities made available to him as a Vacation Consultant by the United Kingdom Atomic Energy Research Establishment, Harwell, during this work and is also grateful that his contribution to the research reported has been sponsored in part by the Office, Chief of Research and Development, U.S. Department of the Army through its European Research Office.

References

CLARKE, F. J. P., and GARTON, W. R. S., 1959, *J. Sci. Instrum.*, **36**, 403.

DOYLE, W. P., 1958, *J. Phys. Chem. Solids*, **4**, 144.

HOLLAND, L., 1956, *Vacuum Deposition of Thin Films* (London: Chapman and Hall), pp. 517–18.

REILING, G. H., and HENSLEY, E. B., 1958, *Phys. Rev.* **112**, 1106.

Notes and comments

Symposium on electron spin resonance

The Chemical Society will hold a Symposium on electron spin resonance at Queen Mary College on Thursday, 2nd November 1961. Further information may be obtained from The Chemical Society, Burlington House, London, W.1.

International Conference on crystal lattice defects

An International Conference on crystal lattice defects will be held in Kyoto, Japan, from 7th–12th September 1962. The Conference is being organized by the Physical Society of Japan. The International Union of Pure and Applied Physics is considering sponsorship of the Conference.

The Conference will cover point defects in metals, semiconductors and ionic crystals, as well as extended defects closely related to point defects. Suggested items of discussion are, for instance, electronic processes in imperfect crystals, properties of single and multiple point defects, radiation damage, association, dissociation and diffusion of point defects. Purely phenomenological studies will not be included, nor work that is ambiguous from the viewpoint of lattice defects.

The attendance will be limited to actually active research workers in the field of lattice defects, and others who may be preparing to enter work in this field. The meeting will be arranged to reserve ample time for discussions. The language used will be English.

All correspondence should be addressed to: Lattice Defects Conference Secretariate, c/o Prof. R. R. Hasiguti, Department of Metallurgy, University of Tokyo, Bunkyo-ku, Tokyo, Japan.

Conference on the physics of semiconductors

The Institute of Physics and The Physical Society, on behalf of the International Union of Pure and Applied Physics and the British National Committee for Physics, is arranging an International Conference on The Physics of Semiconductors which will be held at the University of Exeter from 16th to 20th July 1962. The Conference is planned to follow the previous sequence of Conferences on the physics of semiconductors which were held in Reading in 1950, Amsterdam in 1954, Garmisch in 1956, Rochester in 1958 and Prague in 1960.

Accommodation will be provided in Halls of Residence at the University. Further information regarding the Conference may be obtained from the Administration Assistant, The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1.

Improved electron gun for zone-melting bombardment

by M. COLE, M.A., Ph.D., C. FISHER,* B.A., and I. A. BUCKLOW, M.A., Ph.D., Metals Research Ltd.
91 King Street, Cambridge

MS. received 12*th June* 1961

Abstract

A simple electron focusing device for floating-zone electron-bombardment furnaces is described. The filament is surrounded by and electrically connected to a water-cooled copper cage, which focuses the electron beam some 3 cm vertically below the filament. The anode is thus largely protected from evaporation from the filament, the molten zone is not obscured and evaporation and spluttering from the zone do not contaminate the filament. The sharpness of the focused beam may be altered by varying the filament diameter and position. Rods of between 3 mm and 12 mm diameter have been successfully zoned in a cage of the dimensions given.

1. Introduction

The electron-beam, floating-zone technique is widely used for refining and producing single crystals of refractory or reactive materials, and the advantages of the method are well known. In most of the published work, the electron gun is basically of the type originally used by Calverly (1957) in which horizontal plates, above and below a ring filament, focus the electron beam horizontally inwards on to that part of the anode surrounded by the filament. This arrangement suffers from the following disadvantages:

(i) Spluttering and evaporation of metal from the molten zone contaminates the filament causing loss of emission and possibly failure.

(ii) It is frequently difficult to observe the molten zone because of the bright filament wire in the line of sight.

(iii) Evaporation from the filament contaminates the specimen.

We have developed a simple gun in which the filament is removed and protected from the immediate neighbourhood of the molten zone and in which electrons from the cathode are made to flow in a curved path. This arrangement eliminates the first two difficulties and greatly minimizes the third, without increasing the complexity of the control mechanisms.

2. Procedure adopted in design

The design of electron beam focusing devices is an involved process which is made even more difficult if space charge effects are considered (Beck 1953, Pierce 1954). However, it is possible to obtain, fairly simply, an approximate solution of the electrostatic field distribution if space charge is ignored.

Practical requirements in the design of the gun were (i) that it should be mechanically simple and axially symmetrical, and (ii) that the anode should be shielded from direct line of sight of the cathode; these two factors lead to a structure essentially similar to the final form which is shown in Fig. 1.

An approximate construction for the field distribution showed that the equipotential surfaces bulged away from that part of the anode exposed within the cage; their form suggested that electrons would be focused on the anode and that vertical movement of the bottom plate would alter the position and degree of the focus. Thus the proportions of the structure for optimum focus were obtained by experiment.

The shapes of the electron trajectories have since been investigated (Hardisty 1961, unpublished), and the focus predicted on the basis of the above approximations coincides with that observed. In view of the heavy currents and long electron paths, it is concluded that the presence of ionized gas (metal evaporated from the anode and gas within the

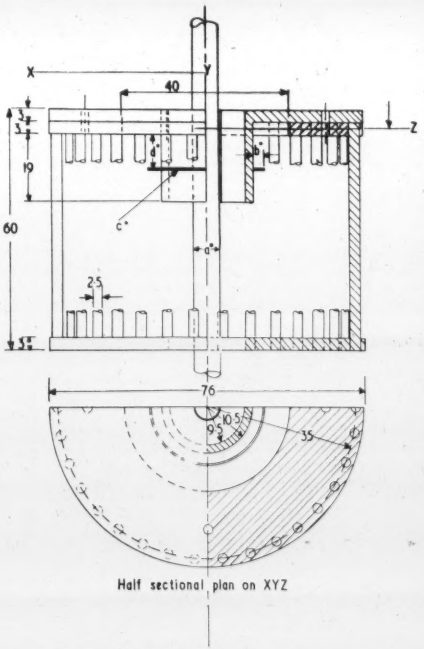


Fig. 1. The focusing-cage electron gun. All dimensions are in millimetres.

$a^* = 6$ optimum (3–12 possible), $b^* = 2{\cdot}5$–$4{\cdot}0$ depending on focus required, $c^* = 0{\cdot}4$–$0{\cdot}8$, $d^* = 8$–14 depending on current required.

* Now at Department of Engineering, Cambridge University.

chamber) is responsible for annulling the space charge effects.

3. Operation of the gun

The device is illustrated in Fig. 1 and consists of a brazed copper cage bolted to a water cooled top plate. The top plate carries a tungsten filament and a copper tube which partially shields the anode from the filament. The whole cage assembly is electrically connected to one end of the filament and the specimen to be melted is made 6–10 kv positive with respect to the cathode assembly; electrons are focused to heat a zone approximately 3 cm below the filament. By adjusting the filament loop diameter it is possible to vary the focus from about 0·5 mm to 6 mm; focusing is also slightly sensitive to the vertical position of the filament which should be as high as possible* in order to reduce contamination of the specimen. Fig. 2 is a photograph of the equipment in operation; the anode is a 6 mm graphite rod on which can be seen a fine, bright line focus in the middle of a zone heated by conduction.

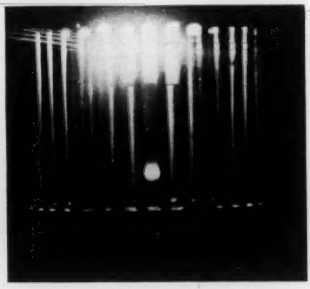


Fig. 2.

When a zone is melted in the gun, any metal blown off by out-gassing emerges in a relatively small solid angle approximately normal to the axis of the rod and thus avoids the filament. An unobstructed view of the zone is obtained and very little of the tungsten vapour evolved from the filament can impinge on the anode.

The device behaves as a diode operating in the saturated regime, therefore beam current is controlled solely by the electron emission from the filament, i.e. by filament temperature. Filament current control requires little power and is simply and cheaply achieved by a magnetic amplifier. Moreover, at any beam current, the filament is run at the lowest possible temperature, thus evaporation and sagging are reduced and filament life is prolonged.

Acknowledgments

The authors wish to thank Dr. C. W. B. Grigson of the Engineering Laboratories, Cambridge, for his valuable assistance in the preparation of this paper.

References

CALVERLEY, A., DAVIS, M., and LEVER, R. F., 1957, *J. Sci. Instrum.*, **34**, 142.

BECK, A. H. W., 1953, *Thermionic Valves*, Chaps. 4, 5, 6 (Cambridge: University Press).

PIERCE, J. R., 1954, *Theory and Design of Electron Beams* (New York: Van Nostrand).

* The maximum current available is progressively reduced as the filament approaches the top plate.

Electrical conductivity and photoconductivity of thin anthracene layers in vacuum

Abstract. In order to eliminate, as much as possible, the surface phenomena produced by air and water vapour adsorption, cells were prepared inside the vacuum system used also for the electrical conductivity and photoconductivity measurements. It was observed that the dark and photocurrents increase with increase of pressure.

The electrical conductivity and photoconductivity of anthracene was studied in air, vacuum and other gaseous atmospheres using this substance in the form of single crystal (Mette and Pick 1953, Chynoweth and Schneider 1954), pressed crystalline powder (Eley *et al.* 1953, Eley and Parfitt 1955) and of a thin layer obtained by the evaporation of the solvent (Vartanyan 1950) or by thermal evaporation in vacuum (Northrop and Simpson 1956, 1958, Gheorghiță-Oancea and Teodorescu 1960). When the study of these phenomena was carried out in vacuum, the photoconductive cell was prepared in air and then introduced into the vacuum.

In order to eliminate, as much as possible, the surface phenomena produced by the air and water vapour adsorption, etc., the cells were prepared inside the vacuum system used also for the electrical conductivity and photoconductivity measurements. The cell was made by thermal evaporation in vacuum on to a quartz support on which the electrodes were painted beforehand.

The installation includes (*a*) the tube in which the cell was prepared by thermal evaporation in high vacuum ($\simeq 10^{-6}$ mm Hg) and where the determinations were made (*b*) the electrical assembly for measurements of dark currents down to 10^{-15} A and (*c*) the optical assembly for the study of the photoconductivity.

The dark current of the support was determined before laying down the substance on the quartz plate, on which were painted the two aquadag electrodes. The operation was carried out in the same subsequent working conditions for pressure, temperature and stress, after the deposition of the substance. In the same physical conditions were measured also the currents, when the cell is illuminated with $\lambda = 3660$ Å.

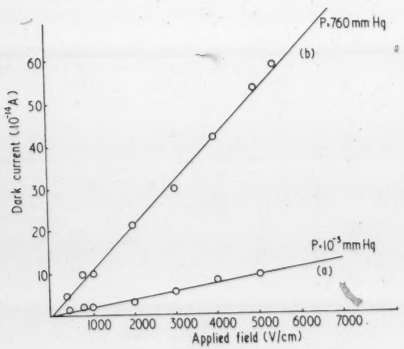


Fig. 1

The measurement of dark currents and photocurrents was made for various fields up to 10000 v/cm for pressures from 10^{-5} mm Hg to atmospheric pressure.

In Fig. 1(*a*) the dark current is plotted against the applied field at 10^{-5} mm Hg and at room temperature for the anthracene cell. Similar results were obtained also for pressures of 10^{-3}, 10^{-2}, 10^{-1}, 1 and 10 mm Hg. For pressures greater than 50 mm Hg the dark current grows with increasing pressure, reaching at 760 mm Hg values five to six times larger than at 10^{-5} mm Hg. In Fig. 1(*b*) the dark current is plotted against the applied field at the atmospheric pressure.

Figure 2(*a*) shows the variation of photocurrents with applied field at 10^{-5} mm Hg and at room temperature. The dependence of the photocurrent on pressure becomes appreciable at about 10^{-2} mm Hg when the photocurrent increases suddenly. In Fig. 2(*b*) the variation of the photocurrent with the applied field is plotted at atmospheric pressure, where the photocurrent is doubled.

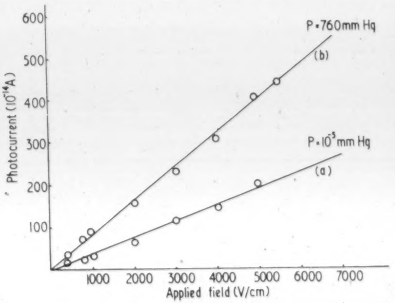


Fig. 2

From Figs 1 and 2 and other determinations, the linearity of the dependence of the dark currents and photocurrents on the applied field is evident up to values of 10000 v/cm; this result agrees with other papers. The linearity is, however, perturbed if the contacts of electrodes are not good or the illumination is not correctly managed; in these cases the curve approaches the vertical axis.

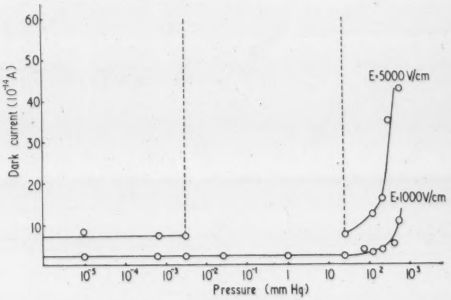


Fig. 3

Figure 3 shows the dependence of the dark current on pressure for two voltages applied to the cell, 100 v and 500 v. In Fig. 4 the same variation is plotted for the photocurrent. We have observed that at a given pressure a sudden increase

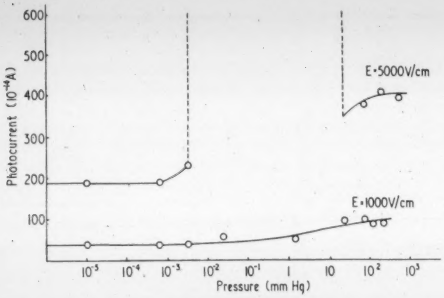


Fig. 4

in the dark current and photocurrent occurs at a particular value of the applied field. The value of these currents is much larger than that marked in the Figs 3(*b*) and 4(*b*) at 500 v and is attributed to the passage of current through the rarefied gaseous medium around the cell.

The experiments connected with this paper show the importance of the surface states which appear owing to the adsorption of air at various pressures. In the case of the anthracene, air adsorption produces an expansion of conductivity and photoconductivity which grows as pressure increases.

We warmly thank Professor I. Agârbiceanu for the attention granted to this work. We also thank Mr. Vasile Böhm for making the electronic equipment and Mr. Emilian Ciolacu for devoted help.

The Physical Laboratory,
The Polytechnic Institute,
Bucharest.

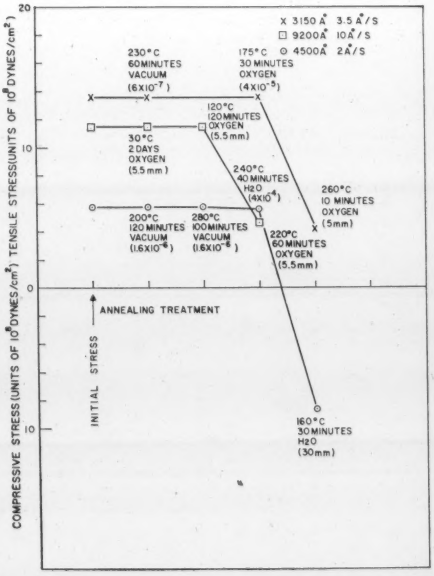
C. GHEORGHIȚĂ-OANCEA
10*th August* 1961

CHYNOWETH, A. G., and SCHNEIDER, W. G., 1954, *J. Chem. Phys.*, **22,** 1021.
ELEY, D. D., and PARFITT, G. D., 1955, *Trans Farad. Soc.*, **51,** 1529.
ELEY, D. D., PARFITT, G. D., PERRY, M. J., and TAYSUM, D. H., 1953, *Trans Farad. Soc.*, **49,** 79.
GHEORGHIȚĂ-OANCEA, C., and TEODORESCU, G., 1960 *Bul. Inst. Polit. Buc.*, **22,** 13.
METTE, H., and PICK, H., 1953, *Zeit-f. Phys.*, **134,** 556.
NORTHROP, D. C., and SIMPSON, O., 1956, *Proc. Roy. Soc.*, A, **234,** 124.
—— 1958, *Proc. Roy. Soc.*, A, **244,** 377.
VARTANYAN, A. T., 1950, *Dokl. Ak. Nauk*, **71,** 641.

Annealing of residual stress in silicon monoxide films

L. Holland, *et al.* (1960), have reported in a correspondence note to this *Journal*, that the internal stresses in silicon monoxide films deposited on to soda lime glass can be relieved if (i) the films are annealed at 200°c or above, or (ii) the temperature of the substrate during deposition is 200° c or above. We have made a study (*Vacuum Symposium Transactions* 1961) of the stresses in silicon monoxide films, and in this note we will describe some ramifications of the annealing process.

Silicon monoxide films were deposited at pressures less than 5×10^{-6} mm Hg on to pure nickel substrates which were clamped at one end to form a cantilevered beam. The film stress was determined from a knowledge of the force required to restore the free end of the beam to its unstressed position (Story and Hoffman 1957). The substrates were maintained at room temperature. The evaporation source consisted of a planar tungsten spiral imbedded in granular SiO. For a source to substrate distance of 8½ in., the deposition rates varied from 4 to 450 Å/s. For source temperatures below 1250° c deposition rates less than 10 Å/s were obtained, and the resultant films were in a highly stressed condition, e.g., a tensile stress of $1{\cdot}6 \times 10^{9}$ dyn/cm^2 at an evaporation rate of 5 Å/s. The results of various annealing processes on three films (initially under a high tensile stress) are shown in the Figure. As seen heat treatments, *in vacuo*, of up to 2 hours at temperatures of 200 to 280° c produced no detectable change in the stress. Similar treatments in atmospheres of O_2 and water vapour at pressures of 4×10^{-5} to 1×10^{-4} mm Hg likewise produced no detectable stress change. However, if the partial pressure of O_2 or water vapour was increased to the mm Hg range,


COMPRESSIVE STRESS(UNITS OF 10^8DYNES/cm^2) TENSILE STRESS(UNITS OF 10^8 DYNES/cm^2)
20
10
0
10
X 3150 Å 3.5 Å/S
□ 9200 Å 10 Å/S
175°C ⊙ 4500 Å 2 Å/S
230°C 60MINUTES VACUUM (6×10^{-7})
30 MINUTES OXYGEN (4×10^{-5})
120°C 120MINUTES OXYGEN (5.5mm)
30°C 2 DAYS OXYGEN (5.5mm)
240°C 40 MINUTES H_2O (4×10^{-4})
260°C 10 MINUTES OXYGEN (5mm)
200°C 120 MINUTES VACUUM (1.6×10^{-6})
280°C 100 MINUTES VACUUM (1.6×10^{-6})
220°C 60 MINUTES OXYGEN (5.5mm)
ANNEALING TREATMENT
INITIAL STRESS
160°C 30 MINUTES H_2O (30mm)

the stress was reduced substantially, and in the case of water vapour at 30 mm Hg the stress actually turned compressive after a 30 minute anneal at 160° C. Although both oxygen and water vapour produce stress relief, water vapour is noticeably more effective and will cause this relief at a somewhat lower temperature.

It is well known that soda lime glass contains large quantities of water which is desorbed in vacuum at elevated temperatures (Siddall 1959, Todd 1955), particularly above 200° C. A possible explanation of Holland's results is that the stress relief occurs when this water from the soda lime glass reacts with the silicon monoxide at elevated temperatures. (The presence of an aluminum film between the substrate and SiO does not exclude the possibility of such a reaction. For example, recent work (Dasilva and White 1961) on aluminum–aluminum oxide capacitors indicates oxygen may diffuse quite rapidly through an evaporated aluminum film.) This is substantiated further by the fact that silicon monoxide films with little or no stress can be prepared by deposition at low rates in water vapour partial pressures above 10^{-6} mm Hg, e.g. zero stress at water vapour partial pressure of 7×10^{-5} mm Hg and a deposition rate of 10 Å/s (source temperature of 1183° C).

International Business Machines Corporation,
Thomas J. Watson Research Center,
Yorktown Heights, New York,
U.S.A.

J. PRIEST
H. L. CASWELL
3rd August 1961

DASILVA, E., and WHITE, P., 1961, *J. Electrochem. Soc.*, in the press.

HOLLAND, L., PUTNER, T., and BALL, R., 1960, *Brit. J. Appl. Phys.*, **11**, 167.

SIDDALL, G., 1959, *Vacuum*, **9**, 274.

STORY, H. S., and HOFFMAN, R. W., 1957, *Proc. Phys. Soc. B*, **70**, 950.

TODD, B. J., 1955, *J. Appl. Phys.*, **26**, 1238.

Vacuum Symposium Transactions, 1961 (New York: Pergamon Press), in the press.

The letter by Priest and Caswell on the nature of stresses in silicon monoxide films is of interest to us. Apart from our work on the removal of stresses in silicon monoxide films on aluminium front surface mirrors, we also briefly investigated the nature of the internal forces in silicon monoxide films deposited on to annealed aluminium foils. This latter work was not published but it showed that the stresses in silicon monoxide films can be either tensile or compressional depending on their mode of growth. We observed that silicon monoxide films deposited slowly (1–2 Å/s) at high pressures ($1–2 \times 10^{-4}$ mm Hg) were under compression whereas films condensed at a pressure of 10^{-5} mm Hg were in tensile stress. The tensile stress of these latter films slowly changed to one of compression after a period of atmospheric exposure.

We have concluded from the foregoing that the silicon monoxide films deposited at a low pressure are incompletely oxidized, as can be observed from their optical properties, whilst films deposited at the high pressures tend more to the composition of silica and contain excess gas sorbed during growth. Desorbed water vapour in the vacuum system may chemically react with the growing deposit so that hydrogen is absorbed in the film. Whilst we do not discount the possibility that water vapour desorbed by glass during baking reacts with the silicon monoxide film, the relief of stresses in silica like films deposited at high pressures is more likely to arise from gas being *desorbed by the deposit*. Further, as our silicon monoxide layers were on aluminium films, one would expect water vapour desorbed from the glass to react mainly with the aluminium rather than with the silicon monoxide.

The silicon monoxide films studied by Priest and Caswell were deposited at lower pressures and generally at faster evaporation rates than that used by us. Such films are generally deficient in oxygen and would be expected to show a tensile stress due to lattice disorder or differential expansion relative to the substrate. The results of Priest and Caswell show that the tensile stress disappears at a deposition rate of 10 Å/s in water vapour at 7×10^{-5} mm Hg. If the evaporation rate was further lowered and the pressure raised they would probably find that the initial stress became compressional.

Vacuum Deposition Research Division,
Edwards High Vacuum Ltd.,
Manor Royal, Crawley,
Sussex.

L. HOLLAND
T. PUTNER
15th August 1961

New books

Essentials of dielectromagnetic engineering. By H. M. SCHLICKE. (London: John Wiley, 1961.) Pp. xxii + 242. Price 76*s*.

The title of this book requires some explanation, and hence a subtitle has been added, equally clumsy, but at least fairly clear: *An introduction to the thinking in and the use of ferrites and high-permittivity dielectrics*. One feels it would have been simpler just to call the book 'The use of ferrites and high-permittivity dielectrics'. This would have been clear and concise. This clumsiness and verbosity pervades the whole book. Although the author is at pains to point out that he has intended to economize in space by giving principles and teaching the use of these by examples, a generous proportion of the book deals with generalities and explanations of the layout and of subject arrangement. The language is precious and forced and the whole tone of the book is unattractive.

The book is designed for engineers and students who wish to understand the possibilities of the use of 'dielectromagnetics'. The first chapter is a sketchy and purely descriptive introduction to electromagnetism with special reference to high μ and high ϵ media. Here, like in other parts of the book, there appear sudden changes from discussion of fundamentals to examples of rather special applications. One supposes that this is what the author means by "intuitive synthesis based, in part at least, on deductive reasoning from principles". This chapter is followed by a somewhat unsystematic description of ferrite and dielectric material properties.

After these two introductory chapters applications form the rest of the subject of the book. Lumped circuit elements, distributed elements and non-linear and other unusual applications are treated in turn. All these chapters give the impression of being somewhat untidy.

For the price of 76*s*. most students and young engineers would expect a more easily readable and more comprehensive book.

The book is very well produced, has fairly adequate illustrations but a poor list of references and an indifferent index.

K. HOSELITZ

Progress in cryogenics. Vol. 3. Edited by K. MENDELSSOHN. (London: Heywood, 1961.) Pp. vii + 173. Price 45*s*.

The remarkable properties of materials at deep low temperatures are bound to find wide applications sooner or later. A few applications seem already assured of wide usage and many others are receiving detailed consideration; it has been seriously suggested, for example, that the practicability of thermonuclear power generation may hinge on the possibility of producing extremely intense magnetic fields using superconducting solenoids. If temperatures below 60° K are to be used by the non-specialist in complex apparatus, considerable progress will have to be made in rendering the means of production more flexible and more fool-proof. In short, a new branch of engineering must be developed. The editor of the series *Progress in cryogenics* intends to "provide summarizing articles on the whole field of low temperature methods as distinguished from low temperature physics or chemistry". The publication seems well set, therefore, to assist in this new development. The present volume contains six articles.

The first, by Dr. A. J. Croft, deals with the proven methods of production of liquid helium. This is a field in which the viewpoints of physicist and engineer have recently been combined, leading to the marketing of three proprietary liquefiers. As a result research into the properties of materials at low temperatures has been accelerated and, the first step towards the widespread use of deep low temperatures having been successfully negotiated, there is a prospect of bigger things to come. The article by Mr. W. E. Gifford is concerned with the design problems arising from two novel refrigeration cycles. Refrigerators based on these cycles have been constructed and have produced liquid helium and they promise considerable advances in flexibility, reliability and ease of maintenance and operation.

The other four articles reflect the widely different aspects that the term cryogenics can assume. Those by Dr. R. P. Hudson and Dr. C. D. Jeffries deal with methods of nuclear orientation. The temperatures involved here are generally well below 2° K and at this end of the scale the principles of the methods are not easily distinguished from the detailed physics of the working substances. Mr. A. G. Lenfestey's article on heat exchangers deals mainly with the very large installations required for plant producing liquid air or oxygen on a 'tonnage' scale, and he includes photographs of exchangers whose size will surprise anyone familiar only with laboratory liquefiers. Dr. I. E. Smith gives an account of the working of a rocket engine and compares liquid oxygen and liquid hydrogen with other fuels in the matters of intrinsic efficiency and ease of storage and distribution.

The author is in each case very well qualified by experience to write on his subject. If the volume is to be criticized, it must be on the grounds that only the first two articles above are predominantly about cryogenic methods, for nuclear orientation is not likely for some time to concern those whose main interests are in wide applications of low-temperature phenomena, and the last two articles deal mainly with methods of handling of liquid oxygen which may not be simply extrapolated to lower temperatures. This nevertheless reflects fairly well the present situation. To anyone who has gained an interest in the technical problems of low temperatures, the prospect of following the development of the subject in this series is very welcome.

D. H. MARTIN

Non-crystalline solids. Edited by V. D. FRÉCHETTE. (New York, London: John Wiley, 1960.) Pp. xviii + 536. Price 120*s*.

This book consists of papers presented at the Conference on Non-crystalline Solids at Alfred, New York, from 3rd to 5th September 1958. It was sponsored by the National Academy of Sciences, National Research Council and Air Force Office of Scientific Research, Air Research and Development Command. This Conference brought together leading investigators from several fields interested in the structure and properties of non-crystalline solids. The book deals with the scattering of radiation by non-crystalline materials, electronic structure, relaxation phenomena and structure and properties of special systems.

Similitude physique—exemples d'applications à la méchanique des fluides. By A. MARTINOT LAGARDE. (Paris: Gauthier-Villars, 1960.) Pp. 70. Price 14.00 N.F.

Fizichesky Entsiklopedichesky Slovar. Vol. 1. A–D. (Moscow: State Scientific Publishing House, 1960.) Pp. 664. Price 45*s*. Obtainable, in Russian, through Collet's, 44 and 45 Museum Street, London, W.C.1.

Nuclear reactor containment buildings and pressure vessels—Proceedings of a symposium organized by the Department of Mechanical, Civil and Chemical Engineering of the Royal College of Science and Technology, Glasgow, 17th to 20th May 1960. (London: Butterworth, 1960.) Pp. viii + 572. Price 100*s*.

This volume contains 22 papers and discussions devoted to the analysis, design, construction and testing of containment buildings and pressure vessels of the type used in nuclear power plants. It is divided into six sections: Current practice and future trends; Design studies and methods of stress analysis; Shell research—analysis and experiment; Engineering design, fabrication, erection and testing; Open topic contributions; Summaries of the proceedings.

Mathematical handbook for scientists and engineers. By G. A. KORN and T. M. KORN. (New York, Toronto, London: McGraw-Hill, 1961.) Pp. xiv + 943. Price £7 15*s*. 0*d*.

The handbook is a comprehensive reference collection of mathematical definitions, theorems and formulae set out in increasing order of complexity, ranging from elementary algebra to elliptic functions, Legendre functions and so on. The most important formulae are clearly displayed and more detailed discussion is set out in smaller type. The book carries a number of appendices including a comprehensive table of indefinite and definite integrals and numerical tables of most of the more important functions.

The writers hope that the method in which the book is set out will enable readers to obtain a connected survey of mathematical methods.

Magneto-fluid dynamics. Edited by F. N. FRENKIEL and W. R. SEARS. Proceedings of a Symposium Sponsored by the International Union of Theoretical and Applied Mechanics in Co-operation with the National Academy of Sciences–National Research Council. (Washington: National Academy of Sciences–National Research Council, 1960.) Pp. 696 to 1032 + 693, 694 and 695 (the same page numbers are used as in the Reviews of Modern Physics in which these papers also appear). Price $4.00.

This is a collection of papers presented at the symposium, which have already been published in Reviews of Modern Physics, and the page numbers are those used in this publication. The first session included an introductory survey of fluid dynamics and plasma physics, and the following sessions elaborated on these subjects. About fifty papers are included in the volume.

Physical gas dynamics. By A. S. PREDVODITELEV. (Oxford, London, New York, Paris: Pergamon Press, 1961.) Pp. 183 + 3. Price 50*s*.

This volume, translated from the Russian, contains the proceedings of a symposium devoted to work carried out in the Power Engineering Institute of the Academy of Sciences of the U.S.S.R. It deals with the problems of gas dynamics and the thermodynamic properties of air at high temperatures (up to 12000° K) in a wide range of pressures from 0·001 to 1000 atmospheres. Methods of calculating a normal shock with allowance for dissociation and ionization of the air are given. Part of the symposium is devoted to an investigation of the hydrodynamic phenomena accompanying electrical discharges in water.

This book will be of interest to scientific workers, teachers, engineers, and students of senior courses in higher institutes of learning who are specializing in the field of gas dynamics and combustion physics.

Technische Strömungslehre. By BRUNO ECK. (Berlin, Göttingen, Heidelberg: Springer-Verlag, 1961.) Pp. 453 + x. Price DM 31.50.

This sixth edition includes in particular new work on gas dynamics and the motion of solid bodies in fluids.

An introduction to astronomy. By R. H. BAKER. (U.S.A., Canada: D. Van Nostrand Co., 1961.) Pp. 364 + 8. Price 41*s*. 6*d*.

This sixth edition includes information gleaned during the recent International Geophysical Year, and a description of recent developments from radio astronomy.

Cobalt Monograph. Edited by Centre d'information du Cobalt. (Brussels: M. Weissenbruch, 1960.) Pp. 515 + xv. Price B.F.750

This book is a complete survey of cobalt, its occurrence, mining, extraction, the properties of cobalt and its alloys and a review of both metallic and non-metallic applications.

Medical aspects are also covered.

Nuclear instruments and methods. Edited by K. SIEGBAHN. (Amsterdam: North-Holland Publishing Company, 1961.) Pp. 253 + x. Price 90*s*.

This is a report of the proceedings of the Second Accelerator Conference held in Amsterdam in 1960. The Conference was divided into sessions which considered Nuclear physics, Particle research and Physics of radiation research.

Modern physics for the engineer. Edited by L. N. RIDENOUR and W. A. NIERENBERG. (New York, Toronto, London: McGraw-Hill, 1961.) Pp. xv + 383. Price 74*s*.

Notes and comments

Notes for the Preparation of Papers for the *British Journal of Applied Physics*

The following instructions refer to the main points which have to be considered in the preparation of a paper for publication, and authors offering papers to the *British Journal of Applied Physics* for publication are asked to conform to these recommendations.

Manuscripts

Manuscripts should be typed in *double* spacing on paper not wider than 8 in. and not longer than 13 in. Only one side of the paper should be used, and a margin of 1–1½ in. should be left. As alterations in the text cannot be allowed once the paper is set up in type, authors should aim at absolute clarity of meaning and of typing, and should check the typescript carefully before submission. All manuscripts should be submitted in duplicate, and in addition to the fair copies, a set of small copies or prints of the diagrams (not larger than foolscap) must be attached to each MS. (This enables MSS. to be sent to two referees at the same time and so assists rapid publication.)

Abstract. An abstract is printed at the beginning of the paper immediately after the title, name(s) of author(s), and place of employment of author(s). Two extra copies of the 'title and abstract' page are required for the records.

Mathematics. It is not necessary to give detailed derivations of mathematical expressions and formulae in a published paper when the work is straightforward; it is quite sufficient to indicate the method of treatment and the final results.

References. In the text bibliographical references are made by giving the name of the author and the year of publication in brackets, e.g. (Jones 1942), and details are given in the last section, 'References', where the references are arranged in alphabetical order of authors' names and in date order for each author.

Drawings

Drawings should be in Indian ink on tracing cloth, tracing paper or white card, with lettering in soft or blue pencil; lettering of a size suitable for reduction will be inserted by our draughtsman. The drawings should in general be sufficiently large to allow of reduction in printing, and the lines should therefore be bold; the frame lines of graphs should be slightly finer than those of the plotted curves.

Full details are available in the Institute and Society's 'Notes for Authors', obtainable from the Editor and Deputy Secretary, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London S.W.7.

Papers for publication in the *British Journal of Applied Physics* should be sent to the Editor and Deputy Secretary at the same address.

Letters to the Editor

We should like to draw readers' attention to the publication of Letters to the Editor in this journal. We have arranged with our printers to publish letters received up to six weeks before the date of publication, provided they are reported upon favourably by our referees. For example, items for publication in the December issue of the *Journal* should be received in the office not later than 20th October. Letters from the Continent will be published in French or German if necessary to avoid further delay. This also applies to the other two journals of The Institution of Physics and The Physical Society, i.e. the *Proceedings of the Physical Society* and the *Journal of Scientific Instruments*.

Corrigendum

The structure of non-destructive testing by K. W. ANDREWS, *Brit. J. Appl. Phys.*, **12**, 127.

The sentence beginning 'His definition of magnetic methods . . .' (line 18 of introduction) should read 'His definition of magnetic methods was already too narrow and metallurgists and others have been aware of a number of electrical and magnetic methods which have been available for several years (Lewis 1951, Institute of Metals 1958a).

Journal of Scientific Instruments

Contents of the October issue

PAPERS

Instrument for the continuous determination of the mechanical internal energy losses in high-polymer solids over a wide range of temperature frequency. By P. Lord and R. E. Wetton.

Digitizer-to-tape punch-coupling unit. By L. Molyneux, E. E. Schneider, and G. R. Sharp.

Crystal-pulling furnace for gallium arsenide. By R. Wickham.

Method of linearizing thermistor thermometer data in calorimetry. By H. V. Larson, T. T. Myers and W. H. Le Blanc.

Improved microfurnace power supply. By J. H. Welch.

Systematic operational testing of the von Eller optical synthesis machine. By J. C. Coppola.

Instrument for cutting animal tissues into small pieces (< 3 mm³) with preservation of viable cells. By A. L. Sims.

Anti-distortion mountings for instruments and apparatus. By R. V. Jones.

LABORATORY AND WORKSHOP NOTES

On the adjusting of the de Wolff type Guinier camera. By L. Zsoldos.

Simple storage method for thin self-supporting carbon films. By M. Nobes.

Control of slow motion devices at motor speeds near the stalling value. By W. R. George.

Stabilized power supply for microcalorimetry. By J. D. Hill and E. A. Steer.

Flexible coupling and overload safety device. By F. Deutsch.

CORRESPONDENCE

Simple electromanometer utilizing the CdS photoconductive cell. From T. Azuma, T. Kanno and T. Hirota.

Monitors for diffractometer X-ray beams. By K. D. Chandrasekaran.

NOTES AND NEWS

New instruments, materials and tools — Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.

REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611-2.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12s. 6d. each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.

CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

PRINTED IN GREAT BRITAIN BY UNWIN BROTHERS LIMITED, WOKING AND LONDON

Conference on Electron Microscopy—Nottingham, July 1961

Abstract

The annual conference of the Electron Microscopy Group of the Institute of Physics and the Physical Society was held in the Department of Chemistry, the University of Nottingham, from 10th to 14th July 1961. A general account of the proceedings is given.

PROFESSOR D. D. Eley (University of Nottingham) opened the conference by welcoming the two hundred and sixty participants and explaining the workings of the magnificent new lecture theatre.

Metal physics

In an invited lecture Professor L. F. Bates (University of Nottingham) described the domain structures of various magnetic materials as seen in the light and electron microscopes. Domain structures in single crystals and polycrystals, the relationship between domain spacing and applied field and the domain patterns near imperfections and inclusions, were then treated in detail. This clear and authoritative exposition ended with a plea for the more extensive use of the electron microscope, particularly in deriving quantitative information. This theme was continued by Dr. D. J. Craik (Department of Physics, University of Nottingham) who described results of an electron microscope study of the domain structure in fine grained materials and compacts. The critical particle size, at which splitting into domains gave no reduction in energy, was estimated from the limiting size of particles split into two domains and this value agreed well with calculated radii. Dr. J. Silcox (Cavendish Laboratory, University of Cambridge) and Dr. P. Gaunt (University of Sheffield) had studied thin foils of an aged Au–1·5% Co alloy. Most of the ferromagnetic cobalt precipitates appeared as disks parallel to [100], the disks being arranged in sheets also on [100]. The shape anisotropy of these particles was thought to make a small contribution to the coercivity of the alloy. With thin foils of nickel it was possible to view dislocations and domain walls simultaneously. Clearly the electron microscope could contribute much to improve our knowledge of magnetic structures. Mr. P. L. Ryder (Department of Metallurgy, University of Leeds) described the preparation of thin foils from Calder quality, commercial uranium strip. Although the metal oxidized very readily during electropolishing, there were small, oxide-free regions suitable for transmission microscopy and in such areas dislocation networks were observed with good contrast and resolution. In discussion of this paper Mr. B. Hudson, Dr. K. H. Westmacott and Dr. M. J. Makin (Atomic Energy Research Establishment, Harwell) pointed out that the surface film on thin foils of uranium could be removed in some 20 seconds by a solution of 75 cm^3 sulphuric acid, 7 cm^3 water and 18 cm^3 glycerol, used at 6 v open circuit with a Pt or Ni cathode. The neutron irradiation of α-uranium had been studied using thin foils cleaned in this manner.

The study of iron powders using surface replicas and microtome sections was described by Dr. W. I. Mitchell (International Nickel Co. (Mond) Ltd., Birmingham). The powder was formed by decomposing iron pentacarbonyl and showed a characteristic ring structure. Techniques for studying stress corrosion in copper alloys were reported by Mr. D. Tromans (Department of Metallurgy, University of Leeds). Carbon replicas from the surface of massive specimens of 70/30 brass and thin, ultramicrotome sections were used to examine the cracks formed during stress corrosion in ammonia vapour, whilst stress corrosion of thin foils prepared by electropolishing was also investigated. The role of stacking faults in the stress corrosion process was discussed. Mrs. H. M. Scott (National Engineering Laboratories, Glasgow) had studied the behaviour and mechanism of rupture of oxide films on aluminium during deformation. The results were related to the frictional properties of the metal and their significance with respect to exo-electron emission from strained metal faces was also discussed. Mr. H. McArthur and Mr. D. G. Teer (English Electric Co. Ltd., Leicester) described an electron microscope study of the mode of adhesion of nitrided steel surfaces in static contact at high temperatures.

Dr. D. W. Pashley and Mr. A. E. B. Presland (Tube Investments Research Laboratories, Hinxton Hall) discussed the relation between specimen contamination and the movement of dislocations produced in metal films during electron microscope examination, illustrating the effects with a short cine-film. The rate of contamination was controlled by surrounding the specimen with a relatively high pressure of oxygen and experiments with evaporated single crystal films of gold showed that buckling and dislocation movement arose entirely from the formation of the contamination layer and were independent of the cooling effect of the gas. A transmission electron microscope technique for measuring the dihedral angle at twin/grain boundary intersections was described by Dr. M. C. Inman (National Physical Laboratory, Teddington). The twin boundary energy could be determined from these measurements and for Cu, a value of 11 erg cm^{-2} was found, which gave an estimate of 22 erg cm^{-2} for the stacking fault energy in this metal. The addition of 1 wt % Sb to Cu reduced the surface, grain boundary and twin boundary energy by a factor of two and this was interpreted as evidence for segregation to stacking faults. Mr. G. Pollard and Mr. R. A. Sinclair (Department of Metallurgy, University of Leeds) reported a transmission electron microscope study of thin foils prepared from aged Al–4% Cu and O.F.H.C. copper sheet crept near $0{\cdot}5\,T_M$. Inter-dislocation and dislocation-precipitate reactions, leading to the formation of sub-grain boundaries, were observed, whilst during creep the θ' precipitates in the aged alloy were sheared. Mr. W. K. Armitage and Dr. P. M. Kelly (Department of Metallurgy, University of Leeds) had examined thin foils of a commercial Cu–1·8% Be alloy, aged below 325° C. The hardening mechanism at these ageing temperatures involved the formation of ordered regions of the CuAu I type in the face centred cubic α-phase whilst an atomic mechanism was proposed for the deformation of the ordered alloy by twinning rather than slip. Dr. P. M. Kelly also described his work on thin foils of lower bainite in 0·65% C low alloy steel. The precipitated carbide was found to be cementite with the same arrangement and orientation relationship as the carbide in tempered high carbon steels. The implications of the similarity between

lower bainite and tempered high carbon martensite were also discussed. A study of the stacking fault energy of some stainless steels and Cu–Ni alloys was reported by Mr. D. Dulieu and Mr. D. Tromans (Department of Metallurgy, University of Leeds). Both in a Cu–20% Ni alloy and in 'Kunifer 5', cross-slip was observed, indicating a high stacking fault energy. In 18/10 stainless steels, the stacking fault energy decreased with nitrogen content and in the high nitrogen steel there was some evidence for the segregation of nitrogen to stacking faults. The use of the electron microscope in measuring the equilibrium groove angles developed at the junction of grain boundaries with the free surface of copper and Cu–Sb alloy wires was described by Mr. H. R. Tipler (National Physical Laboratory, Teddington). Using previously determined values of the surface energy, the grain boundary energy could be calculated and it was found that the addition of $\frac{1}{2}$ to 1 wt% of Sb reduced the grain boundary energy by about 50%. The change of surface energy with crystallographic orientation was being investigated with a goniometer stage (see later). The final paper in this session was given by Mr. C. Baker and Mr. G. W. Groves (Department of Metallurgy, University of Cambridge). Thin foils of non-metallic materials were prepared by the repeated cleavage of single crystals, subsequent thinning by chemical polishing being used when necessary. The stacking fault energy of graphite was deduced from the width of the extended dislocations and the values compared with those calculated on the basis of isotropic and anisotropic elasticity. The annealing of dislocation loops in MgO and in graphite was studied and the activation energy of self-diffusion in graphite was found to be 6·5 ev.

Electron diffraction

In an invited address Dr. T. B. Rymer (Department of Physics, University of Reading) discussed recent developments in the theory and practice of electron diffraction and interference. To extend the scope of these techniques in structure analysis, the measurement and interpretation of the observed intensities still needed a firmer theoretical basis, the major obstacle being the calculation of the amplitude of the electron wave field in the structure itself, as distinct from that of the primary beam. The simple kinematic theory was outlined, followed by the alternative approaches offered by the 2-wave approximation of the dynamical theory and by refining the kinematic theory. Japanese work (Fujimoto 1959) over the last few years had used matrix methods to give an explicit formulation of the intensities of the scattered wave in the form of a power series. The intensities can thus be calculated to any degree of approximation; taking only the first few terms of this series, reasonable agreement with previous theories is obtained, though the significance of certain numerical factors is still uncertain. If more terms are included, difficulties arise, especially when dealing with thicker crystals at lower electron speeds. Current trends in electron interferometry were then outlined. Marton, Simpson and Suddeth's (1959) arrangement of a series of carefully oriented crystals to provide interfering beams is simple and is independent of electron speed, though the difficulty in obtaining perfect crystalline films is delaying its development as an interference microscope. The bi-prism lens developed by Möllenstedt and Duker (1956) was described and the inner potentials of crystal structures may now be determined in elegant fashion. Finally the work of Jönsson (1961) in constructing a copper grating for electron diffraction, with slits 1 μ wide, spaced 2 μ apart, was mentioned.

Mr. J. C. North and Mr. J. C. Henderson (Post Office Research Station, Dollis Hill) had modified the stereo-holder in the Elmiskop I so that a sample 3 mm^2 and $\frac{1}{4}$ mm thick could be placed in the microscope for reflection diffraction, without removing the objective pole piece. The sample could be positioned by the normal stage controls and tilted to give the best diffraction pattern, the circularity of the rings being quite good. Mrs. H. M. Scott (National Engineering Laboratories, Glasgow) discussed the abrasive particles present in the oils of internal combustion engines and which are an important cause of wear. Their size and shape had been determined by electron microscopy and the majority had been identified chemically on the basis of their characteristic diffraction patterns. Reflection electron diffraction studies of the surfaces of graphitic and non-graphitic carbon developed by unidirectional rubbing were described by Dr. J. W. Midgeley and Mr. D. G. Teer (English Electric Co. Ltd., Leicester). Distinct orientations were observed and these were interpreted in terms of the basal plane slipping of the plate-like crystallites, the preferential removal of such crystallites being only of minor importance. The correlation of these results with the frictional characteristics of the surfaces suggests that the low friction of graphitic carbon is due mainly to the preferred orientation of the crystallites. Dr. D. W. Pashley and Mr. A. E. B. Presland stated that thin foils of polycrystalline CuAu had been ordered by annealing at 350° c to give the tetragonal CuAuI structure, transmission specimens having been prepared by electropolishing. A fine scale twinned structure results from the ordering process, the crystallography of which has been carefully analysed by comparing the micrographs and selected area diffraction patterns. Many examples fit well with the expected [101] type twin planes; cases of poor fit can be explained in terms of [101] twinning, if allowance is made for the curling which sometimes occurs at the edge of electropolished foils. Mrs. A. B. Franklin and Dr. D. W. Pashley (Tube Investments Research Laboratories, Saffron Walden) had evaporated very thin layers of face centred cubic metals on to smooth [111], single crystal surfaces of gold and silver, inside an electron diffraction camera. Since the rate of deposition was low, the early stages of alloying could be studied. The reflection diffraction patterns indicated that for certain substrate–deposit combinations, alloying occurs in the early stages due to the penetration of deposit atoms with high thermal energy into the substrate, at substrate temperatures which normally would not allow any appreciable interdiffusion. Where there is only limited solubility of the deposit metal in the substrate, the surface layers of the latter become supersaturated with deposit metal which can later diffuse out in annealing. Dr. D. W. Pashley and Dr. M. J. Stowell (Tube Investments Research Laboratories, Saffron Walden) described diffraction patterns from evaporated films of gold and silver prepared in [100] and [111] orientations and containing many twinned structures. Besides diffraction spots from the twins, certain anomalous additional spots occur which apparently arise from double diffraction effects. With randomly orientated, polycrystalline material, the additional spots introduce extra rings, all of which can be accounted for in terms of the double diffraction effect. The extra rings, however, are not consistent with the formation of small amounts of hexagonal close-packed phases, as other workers have suggested.

Instruments and attachments

Dr. P. Duncumb (Tube Investments Research Laboratories, Hinxton Hall) reported that when a low d.c. voltage was

applied across a new type of ceramic-bonded phosphor screen, it became highly sensitive to light, electrons and x-rays. This phenomenon afforded the basis of a simple image intensifier and preliminary details of its application in electron microscopy were given. The results were promising but the overall gain needed to be improved, so that half-tone effects could be registered. The present rather long delay time of the device might be turned to advantage in producing image storing systems. Mr. C. Washington and Mr. R. W. Fearnhead (London Hospital Medical College) discussed the unidirectional tilting of specimens through large angles and described a device capable of $\pm 65°$ tilt. It had proved simple and reliable in use, the adjustments in focusing and stage controls being quickly learned. Much interest was shown in this device and the great improvement in contrast on tilting would prove useful, especially with biological structures. Dr. E. H. Andrews (The Natural Rubber Producers Research Association) stated that no straining devices for the electron microscope capable of applying very large strains were available. Natural rubber and other elastomers, however, undergo partial crystallization changes at strains in the range 200–400% and to investigate these a specimen capsule providing extensions up to 1000% had been designed. A pair of sliding jaws, initially separated by 0·1 mm, can be opened using a lever system which is push-button operated, to give a gap of 1 mm. The restoring force is provided by a small rubber band encircling the jaws. Successful studies of strain-induced crystallization were illustrated by slides. To tilt specimens through large angles and to avoid shift of the visible field when inducing tilt, Mr. D. C. Barnes and Mr. E. Warner (National Physical Laboratory, Teddington) had designed a goniometer stage based on the principle of locating a ball, in which specimens are mounted in an inverted cone. In their instrument a hemisphere, $\frac{5}{32}$ in. diameter, is used and the push pins which contact the diametral plane are moved to tilt the specimen through $\pm 22°$ about each of the mutually perpendicular axes. Thus a composite angle of about $\pm 30°$ specimen tilt may be obtained. The speakers were congratulated on the precision and accuracy of the goniometer and they stated that a simpler version was being developed. Dr. U. Valdre (Instituto di Fisica, Bologna) also described a goniometer stage with restricted specimen movement and with facilities for cooling the specimen under examination. Dr. K. F. Hale and Dr. D. McLean (National Physical Laboratory, Teddington) had used the new goniometer in an investigation of the ageing of Fe–C and Fe–N alloys. The thin foil specimens were tilted about conventional orientations so that the crystallography of the precipitates with respect to the matrix could be simplified and the growth planes determined.

Dr. A. Strojnik (University of Ljubljana) described the construction of simple electron microscopes with reasonably high resolution. An analysis of the world literature of electron microscopy had shown the need for a simple instrument capable of resolving 30 Å. The completed design was discussed. The lenses were combined in one unit, leading to great mechanical stability, a small number of vacuum connections and extreme compactness. Fixed magnifications of two, five and ten thousand times were decided upon. For the high voltage supply a 50 c/s system was preferred, whilst low drift saturated diodes were used as voltage regulators for the lens systems. Such diodes had excellent long-term stability after the initial warming-up period. A simple combination of magnetic and electrostatic control of astigmatism was used, capable of rapid and certain adjustment, and the final instrument was able to resolve 25 Å regularly. This paper had an enthusiastic reception and in discussion Dr. Strojnik gave details of a new stage control for moving the specimen with extreme accuracy. Dr. le Poole (Technical Physics Laboratory, Delft) congratulated the speaker in achieving such an efficient and reasonably priced instrument, but nevertheless pleaded for work on high resolution microscopes to continue, since for certain classes of work they were indispensable.

This session was completed by two invited lectures, both of which were received with great enthusiasm. Mr. T. Mulvey (Associated Electrical Industries Ltd., Aldermaston) gave a fascinating account, enlivened by many entertaining asides, of the development of the electron microscope. He had been impressed by the lack of agreement amongst textbook writers on this subject, by the hesitant acceptance that electrons had optical properties and by the fact that possibly in no other field of applied science had so many prophecies later been proved false. The instrument had developed, not as the brain-child of the research physicist but rather from electrical engineers using the cathode-ray oscillograph as a means of studying surges in high voltage transmission lines. It was striking that the dual nature of cathode rays, waves *and* particles, had been recognized well before 1900 whilst the concentrating effect of a solenoid on such rays had been known before 1890. The cathode-ray oscillograph was gradually improved; plain tungsten filaments and oxide-coated ones were developed alongside the known cold cathode systems; photographic recording within the vacuum of the oscillograph was also perfected. As early as 1924 Gabor had effectively increased the concentrating effect of the coils by encasing them in soft iron sheeting whilst in 1926 Busch had demonstrated that for an inhomogeneous magnetic field of rotational symmetry produced by a short solenoid, the relationships involving focal length, field strength and electron speed fitted well-known light optical formulae. Even so, he did not realize that he was dealing with an optical system and made no suggestion that his 'lens' could be used in a microscope. Thus it was as late as 1935 before it was realized that electron optics were possible and Rudenberg is now credited with the first invention of an electron microscope as such, the first instruments being essentially cathode-ray oscillographs using techniques which had been known for over thirty years. Even in its early stages, many opined that it would be of limited application only, suitable for robust specimens but not for organic materials. By 1939, however, the situation had improved and for the first time the resolution of the electron microscope exceeded that of the light instrument, and from then onwards, with various setbacks due to the last war, the development of electron optics has been steady and impressive. In this account it is impossible to do justice to the many pioneers mentioned by Mr. Mulvey, but clearly his audience looked forward to reading a full, published account of his historical researches.

In the second lecture Dr. V. E. Cosslett (Cavendish Laboratory, University of Cambridge) said in view of the previous remarks, he had some hesitancy in discussing the future trends in electron microscopy. Nevertheless, he maintained that some problems were certain to attract attention within the next five years. In examining very thin specimens at 100 kv with magnetic lenses, the present resolution limit was some 5 or 6 Å. Theory indicated that this performance could be bettered by a factor of two at least, hence such work needed to be done, especially since all adventitious effects had been removed from the electrostatic lens instrument. Again, microscopes with high voltage beams would assume a greater importance. Not only might they offer a better

design compromise between practical and adventitious factors, but they would make possible the examination of material in bulk. Already for metallic and other inorganic structures voltages of 200–750 kv had shown great promise, but for organic materials the indications were that voltages of the order 2000 kv would be necessary. Thus many biologists consider that the problems of radiation damage to the specimens will need to be studied. It was likely that Dr. Cosslett's own laboratory would enter this high voltage field, indeed, work had already started using a commercially available 600–800 kv supply. Another trend would be the chemical analysis of specimens within the electron microscope, using small area diffraction techniques and characterizing the x-rays from selected regions. Already this was possible on a semi-micron scale and this form of analysis was clearly ripe for development. With the acknowledged success of electron microscopy in the research field, one could also expect its extended application, particularly by industry, to problems of a more routine control character, whilst in conclusion the speaker mentioned the plans being considered by the British Joint Committee for Electron Microscopy in training, at all levels, an adequate supply of electron microscopists.

Techniques

The majority of the papers in this session concerned the ion bombardment of specimens, a technique which is likely to become most valuable, particularly in the study of metals. Mr. P. Bowden (Department of Metallurgy, University of Cambridge) had bombarded thin Au foils, prior to microscopic examination, with argon ions of 80 ev energy at $-50°$ c. Even with such low energy beams, contrast changes were observed, due to small dislocation loops which at first line up in rows along the [111] directions. With higher ion doses, long parallel pairs of dislocations are formed. A carbon film on the foil surface appears to reduce these effects. Mr. A. J. Baker (Department of Metallurgy, University of Cambridge) using a similar technique, reported effects with Pt foils. Many discrete dislocation loops were produced, lying along the [111] directions as in Au, and later these joined to form long, single loops. On annealing at 1200° c using a hot stage in the microscope, the loops climbed out of the surface and formed regular networks with individual dislocations lying along [211] directions. In discussion Dr. D. W. Pashley said he had observed similar phenomena due to ion bombardment in the microscope itself, while it was generally agreed that the mechanism whereby such large numbers of one type of dislocation are formed differs from that involved in α-particle and neutron irradiated foils. Mr. A. D. G. Stewart (Engineering Laboratories, University of Cambridge) outlined the features of a three-lens electrostatic scanning microscope with facilities for bombarding the surface under examination with a focused beam of argon ions. To allow bombardment simultaneously with observation, the image was formed by collecting only the high energy reflected electrons. Many photographs, with remarkable depth of field, were shown of etch pit formation and preferred orientation effects in polycrystalline aluminium, whilst in Sn a series of spike formations were observed. Apparently dust particles on the metal surface could shield the underlying layers from ion bombardment. Because of the elegant technique which it involved this paper attracted much discussion. Mr. Stewart said that he hoped to extend the study to other ions such as oxygen and also perhaps to record diffraction patterns of the bombarded surfaces. Similarly, Mr. Massini (Messrs. Trüb-Täuber, Switzerland) described an instrument for studying metal surfaces by emission microscopy whilst under bombardment from an ion beam. The temperature of the specimen was controllable between 20° and 2000° c and since the ion beam could also be varied, metal/gas interactions could be studied under a variety of conditions. Such phenomena are rapid and cine-film techniques are essential. In a fascinating film, which drew much applause, Mr. Massini clearly demonstrated the versatility and usefulness of his instrument by showing melting and oxidation effects in various irons and steels. Mr. R. H. Alderson (Associated Electrical Industries (Manchester) Ltd., Manchester) discussed the factors needed for controlled ion-etching effects in metals. The specimen must be in good electrical contact with the cathode; hence to give a conducting mount it was embedded in a resin containing Al powder. To avoid damage to the mounting film, the ion beam should be confined by an aperture, made of stable material with low sputtering rate, e.g. oxidized Al foil, 5×10^{-3} in. thick. During ion etching, a cylindrical auxiliary anode around the specimen is useful in focusing the beam on to the metal surface whilst a high standard of vacuum cleanliness is essential. Dr. J. Dlugosz (The Shirley Institute, Manchester) had applied ion-etching to a study of tyre-cord adhesives containing methacrylate–graft latex and a resorcinol–formaldehyde resin. These adhesives were etched differentially, the latex component being removed faster than the resin. Presumably the resin has the greater adhesion and acts by retaining the latex, which can then bond during later vulcanization, to the vulcanizate.

Two interesting papers on the field ion microscope were presented. Mr. M. J. Southon (Department of Metallurgy, University of Cambridge) described preliminary studies to extend the scope of such an instrument so that more metal point surfaces could be examined in atomic detail. At present, only refractory metals, e.g. tungsten, may be examined since the points are often evaporated under the intense field stresses. Results using Mo and Ni points were illustrated, together with current/field dependence on point radius. Since the intensity on the viewing screen is low there is clearly a need for image intensifiers to be used. Image quality under various conditions was also studied, the contrast being better at lower field strengths, whilst as field strength increased, contrast weakening and finally reversal set in. Mr. M. Wald (Department of Metallurgy, University of Cambridge) described similar work with W and Mo points of 500 Å radius. Grain boundary widths in Mo were found to be of the order 10 Å whilst a study of edge dislocations enabled the geometry of the atoms in the different planes to be found. Irradiation of W points with particles of 5 mev energy during the operation of the microscope caused damage which consisted of a large number of vacancies. The atoms released in this process mostly adhere to the point surface. The ensuing discussions showed that this new approach to metal structure analysis would prove most useful.

Dr. U. Valdre analysed the dislocation movements in metal foils which arise from the conditions of preparation. The specimens were deformed and first polished on both sides to remove slip steps, then protected on one side with varnish and finally thinned from the other side. Dislocation movement during polishing is revealed by single slip traces and after polishing by double slip traces and preliminary results on stainless steel indicated that quantitative assessments are now possible.

Mr. P. H. Harris and Mr. E. L. Thompson (British Nylon Spinners Ltd., Pontypool) mentioned the difficulties of obtaining melt-cast polymer films which are sufficiently thin

for transmission microscopy. An instrument was described whereby 10 μ thick films of nylon, contained in a silicone oil temperature bath, could be compressed into thin films. If required, the sample could also be ejected quickly into cold quenching baths. Phase changes in nylon 66 were used to show the action of the instrument and Mr. B. J. Spit (Delft) showed similar effects in other forms of nylon melts. Mr. M. Krantz and Dr. M. Seal (Engelhardt Industries Inc., New Jersey, U.S.A.) mentioned that cylindrical pellets, 5 mm long, 0·6 mm diameter, containing Pt sintered with various carriers (including carbon) were available from the petroleum refining industry, where they were used as catalysts. The pellets could be inserted between the faces of a carbon arc and had proved cheap and satisfactory in preparing Pt-shadowed carbon replicas for electron microscopy. Ir and Pt/Ir alloy were also available in pellet form. Miss B. J. Perkins and Mr. D. G. Teer (English Electric Co. Ltd., Leicester) reported a novel method for the replication of small curved surfaces such as fine journal bearings or biological fibres. A thin strip of flexible tape, coated with Formvar, was placed on the surface, moistened with chloroform and the Formvar pressed against the surface by means of a needle. The replicating system was then removed and a metal/carbon replica made in the usual manner, excellent and consistent results being obtained by this method. Though replication techniques based on flexible backing materials had some merits, it was thought that with careless use, stretching might occur and so lead to a wrong impression of the dimensions involved. Mr. F. R. Johnston and Dr. J. Sikorski (Textile Physics Laboratory, University of Leeds) stated that for the routine examination of the surface of a wool fibre, the latter should be partially embedded in Formvar which had been softened by warm acetone. The exposed fibre surface was then shadowed with Pt/C, backed with a collodion film and scribed into squares. The latter were placed on specimen grids and exposed to chloroform vapour, which removes backing film, embedding medium and the original fibre. In interpreting such fibre replicas, abnormal effects where the fibre enters the embedding medium must be considered.

Notwithstanding the successful application of ultramicrotomy to the elucidation of fibre structure, there is still need for alternative techniques such as fragmentation. Mr. D. J. Johnson and Dr. J. Sikorski (Textile Physics Laboratory, University of Leeds) said that much of the controversy in this field arose simply from the differing experimental conditions used by various workers. Sometimes a wide range of microfibrils was revealed whilst often no evidence of fibrillation was seen. However, highly reproducible data on the microfibrils could be obtained by ultrasonic irradiation *without* the prior use of other fragmentation methods and to support this view slides of regenerated and native cellulose, keratin and various synthetic fibres were shown. This theme was continued by Mr. R. A. Bentley and Mr. J. Cartwright (Safety in Mines Research Establishment, Sheffield) who described mechanical dispersion methods for isolating collagen fibrils from fibrosed lung tissue. The tissue is first treated in a power driven apparatus designed to produce large viscous shear forces and this is followed by ultrasonic treatment at 40 kc/s. Since chemical treatments are not involved in the extraction process their effects can be studied separately on the isolated collagen.

This session was completed by the showing of the new film in sound and colour made by Associated Electrical Industries Ltd. entitled 'The electron microscope'. This dealt with the development of the instrument and its associated techniques in a popular but nevertheless scientifically accurate manner and was well received. The production and film techniques were of extremely high standard and obviously this film will be much in demand for lectures to popular audiences.

Sectioning

The final session, attended by 150 people, was a symposium on the physics of sectioning and the embedding process. This has been arranged with the support of member societies of the British Joint Committee for Electron Microscopy with the aim of bringing together the many users of this important technique. In opening the discussion, Mr. R. Phillips (Aeon Laboratories, Egham) said that little attention had been paid to the physical principles involved in cutting ultra-thin sections for electron microscopy. Much of the theory of working metals by machine tools, however, was apparently directly applicable. Contrary to the commonly held view, compression during cutting was relatively unimportant, the main factor being deformation changes due to shear. Where well-defined slip directions were present in the specimen, shear would occur along them, giving rise to steps of slip blocks in the final section. Depending on the crystalline and rheological nature of the specimen, separation of the blocks might occur also, but because of the severe strains imposed by cutting, some plastic deformation of the blocks would be inevitable with all types of sectioned material. Thus the final section would depend on the nature of the specimen and the direction in which it had been cut, these points being well illustrated in the case of ductile metals by sections of foils which had themselves been obtained by sectioning. The steps of slip blocks, proving that shear had occurred, were clearly seen. Moreover, replication studies of the surfaces of sections mostly revealed regular fine lines showing the margins of the sheared blocks. Hence the question of specimen/knife orientation seems of great importance in obtaining sections and merits further detailed study. The speaker felt that additional factors may be involved since certain sections, particularly those from biological materials, gave no indication of sheared slip blocks and possibly changes in texture due to recrystallization may occur. Dr. A. M. Glauert (Strangeways Laboratory, Cambridge) then gave a lucid and detailed account of the various embedding media which were available. Whilst agreeing that the actual cutting process was important, unless the specimens were properly embedded in a suitable medium, no worthwhile sections could be obtained. Though the methacrylates were uniform in properties, easily available, with good penetration of the specimens and capable of being polymerized to varying degrees of hardness, they had some serious disadvantages in that they were damaged both during polymerization and during examination in the microscope. Though polymerization damage could to some extent be minimized, the mass loss during examination was often as high as 50% and the attendant effects, still only partially understood, detracted from picture quality. The epoxy and polyester resins, however, were less troublesome, showing fewer changes during polymerization and a better retention of fine structure during examination. Dr. Glauert then gave many useful formulations of such resins for various purposes, including the new feature of incorporating a flexibilizer which, by reacting with the resins, could control the degree of hardness finally obtained. The various water-soluble embedding media were also reviewed, including AQUON a new component isolated from commercial forms of epoxy resins. The audience found this paper most instructive and were greatly appre-

ciative of the wealth of information Dr. Glauert had made available. Dr. K. Little (Nuffield Orthopaedic Centre, Oxford) continued this theme. During embedding the specimens might react with the growing polymers, volume changes might occur due to shrinkage whilst the monomers might vary in their ability to wet and to penetrate the structures. Again, difficult points, which later drew much discussion, were the optimal times and temperatures to be used for the polymerization process. Dr. Little preferred a rapid process at low temperatures (by the use of u.v. initiation) with partially polymerized methacrylates so that solidification was achieved quickly without the risk of thermal damage. Even so, she agreed that the embedded material should be given time in which to relieve any local strains. There was also much argument as to whether the embedding material should be removed from the section prior to microscopic examination, to minimize the effects of the electron beam on the specimen, the majority being in favour of its retention. In the light of these considerations Dr. Little reviewed a list of possible embedding media and felt that on balance the methacrylates offered the best compromise, though she pleaded for other polymers to be tested. This paper attracted much discussion, the outcome of which was that however good methacrylates might be as embedding materials, their great sensitivity to electron damage in the microscope limited their usefulness. Dr. J. Sikorski (Textile Physics Laboratory, University of Leeds) mentioned the difficulties in interpreting the images of certain types of thin sections, particularly those of complex biological structures. In the paracortex of wool fibres, for example, the fine detail revealed in thin sections was not easily reconciled with the dimensions obtained from x-ray diffraction studies of the material stained with heavy metals. However, if many atom layers are present in the depth of the section, moire patterns from the arrays of microfibrils are possible. This conjecture was strikingly confirmed by comparing the moire patterns from a simple optical analogue with those actually obtained in micrographs of sections of the paracortical region. Thus with certain biological structures it may be advantageous to photograph the sections at various tilts and at different focal settings, in order to determine whether moire effects are present. Mr. M. G. Dobb and Dr. J. Sikorski (Textile Physics Laboratory, University of Leeds) dealt with the difficult problem of finding selective heavy metal stains which could help differentiate the structures in thin sections of biological material, illustrating their remarks with reference to keratin fibres. Though lead, mercury, osmium and silver appear to stain different regions of the fibres, it is not easy to correlate the sites of attachment to the repeat distances obtained from x-ray diffraction data. Clearly more detailed work is needed, though the present indication is that the matrix between the microfibrils is by no means chemically homogeneous. Uranyl acetate, however, appears to be highly selective in its staining properties. Thus the endocuticle of the wool fibre and certain cells in the medulla region are specifically stained. Moreover, since uranyl acetate is regarded as specific for ribonucleic acid, it is likely that the medulla cells are centres for the synthesis of this material and work is in progress to elucidate this point.

In retrospect the symposium proved most valuable and the large amount of detailed information which was freely exchanged should surely lead to a better understanding of the methods of producing thin sections.

University of Leeds,
Leeds 2.

P. M. KELLY
R. REED

References

FUJIMOTO, F., 1959, *J. Phys. Soc. Japan*, **14**, 1558.

JÖNSSON, C., 1961, *Z. Phys.*, **161**, 454.

MARTON, L., SIMPSON, J. A., and SUDDETH, J. A., 1959, *Rev. Sci. Instrum.*, **25**, 1099.

MÖLLENSTEDT, G., and DUKER, H., 1956, *Z. Phys.*, **145**, 377.

List of papers read at the conference

1. Some problems in magnetism.
   Professor L. F. Bates, University of Nottingham.
2. Domain structures in small crystals and lamellae.
   D. J. Craik, University of Nottingham.
3. Thin foil techniques for examining uranium.
   P. J. Ryder, Department of Metallurgy, University of Leeds.
4. Methods of examining iron powder.
   W. I. Mitchell, International Nickel Co. (Mond) Ltd., Birmingham.
5. New techniques for studying stress corrosion in copper alloys.
   D. Tromans, Department of Metallurgy, University of Leeds.
6. Transmitted electron microscopy of non-metallic thin films.
   C. Baker and G. W. Groves, Department of Metallurgy, University of Cambridge.
7. Electron microscope study of the mechanism of rupture of oxide films on aluminium during deformation.
   H. M. Scott, National Engineering Laboratories, Glasgow.
8. The mechanism of adhesion between nitrided steel surfaces.
   H. McArthur and D. G. Teer, English Electric Co. Ltd., Leicester.
9. The relation between specimen contamination and the movement of dislocations produced in metal films during e.m. examination.
   D. W. Pashley and A. E. B. Presland, Tube Investments Research Laboratories.
10. Coherent twin boundary energy in Cu and Cu/Sb alloys.
    M. C. Inman, National Physical Laboratory, Teddington.
11. Dislocation interactions during creep.
    G. Pollard and R. A. Sinclair, Department of Metallurgy, University of Leeds.
12. Phase changes in Cu/Be alloys.
    W. K. Armitage and P. M. Kelly, Department of Metallurgy, University of Leeds.
13. Phase transformations by shear.
    P. M. Kelly, Department of Metallurgy, University of Leeds.
14. Stacking fault energies of alloys.
    D. Dulieu and D. Tromans, Department of Metallurgy, University of Leeds.
15. Electron microscope observations on magnetic materials.
    J. Silcox, Cavendish Laboratory, Cambridge.
16. Grain boundary determination by the electron microscope.
    H. R. Tipler, National Physical Laboratory, Teddington.

17. Recent developments in the theory and practice of electron diffraction and interference.
T. B. Rymer, University of Reading.
18. Modification of the Elmiskop I stereo-holder for reflection electron diffraction.
J. C. North, G.P.O. Engineering Department, Dollis Hill, London.
19. Study of insoluble contaminations in engine oils by electron microscopy and diffraction.
H. M. Scott, National Engineering Laboratories, Glasgow.
20. Electron diffraction study of rubbed graphitic and non-graphitic carbon.
J. W. Midgley and D. G. Teer, English Electric Co. Ltd., Leicester.
21. Electron diffraction and microscopy of twinned structures in evaporated films of gold and silver.
D. W. Pashley and M. J. Stowell, Tube Investments Research Laboratories.
22. The formation of alloys during the earliest stages of deposition of one metal on another.
A. B. Franklin and D. W. Pashley, Tube Investments Research Laboratories.
23. The study of bulk Cu/Au I with special reference to the analysis of the twinned structure.
D. W. Pashley and A. E. B. Presland, Tube Investments Research Laboratories.
24. The early history of the electron microscope.
T. Mulvey, Associated Electrical Industries Ltd., Aldermaston.
25. Ion etching of mounted specimens.
R. H. Alderson, Associated Electrical Industries (Manchester) Ltd.
26. A solid state image intensifier for electron or x-ray microscopy.
P. Duncomb, Tube Investments Research Laboratories, and P. W. Ranby, Thorn Electrical Industries.
27. A high strain specimen stretching device for the Elmiskop I.
E. H. Andrews, The Natural Rubber Producers Research Association.
28. A new goniometer stage for the Siemens electron microscope.
D. C. Barnes and E. Warner, National Physical Laboratory, Teddington.
29. The application of a new goniometer stage to the thin foil investigation of the ageing of iron–carbon and iron–nitrogen alloys.
K. F. Hale and D. McLean, National Physical Laboratory, Teddington.
30. A specimen tilt device suitable for the Elmiskop I.
C. Washington and R. W. Fearnhead, London Hospital Medical College.
31. The development of simple electron microscopes of high resolving power.
A. Strojnik, Ljubljana.
32. Long term prospects in electron microscopy.
V. E. Cosslett, Cavendish Laboratory, University of Cambridge.
33. Low energy ion bombardment.
P. Bowden, Department of Metallurgy, University of Cambridge.
34. Defects produced in platinum by ion bombardment.
A. J. Baker, Department of Metallurgy, University of Cambridge.
35. Simultaneous observation and etching of surfaces by ion beam bombardment in a scanning electron microscope.
A. D. J. Stewart, Engineering Laboratories, University of Cambridge.
36. Investigation of the mechanism of image formation in the field ion microscope.
M. J. Southon, Department of Metallurgy, University of Cambridge.
37. Defects observed in metals with the aid of the field ion microscope.
M. Wald, Department of Metallurgy, University of Cambridge.
38. Application of ionic etching in the study of tyre-cord adhesion.
J. Dlugosz, The British Rayon Research Association, Manchester.
39. The preparation of thin melt-cast polymer films for electron microscopy.
P. L. Harris and E. L. Thompson, British Nylon Spinners Ltd., Pontypool.
40. A modified platinum–carbon replica technique and an application to the study of diamond powders.
M. Krantz and M. Seal, Engelhard Industries Inc., Newark, U.S.A.
41. A single stage replication of surfaces of wool fibres.
F. R. Johnston and J. Sikorski, Textile Physics Laboratory, University of Leeds.
42. The replication of small internal surfaces.
B. J. Perkins and D. G. Teer, English Electric Co. Ltd., Leicester.
43. Fragmentation techniques in the study of the microfibrillar structure of fibres.
D. J. Johnson and J. Sikorski, Textile Physics Laboratory, University of Leeds.
44. The isolation of collagen fibrils from fibrosed tissue.
R. A. Bentley and J. Cartwright, S.M.R.E., Sheffield.
45. Emission microscopy.
A. Massini, Trüb-Täuber, Switzerland.
46. The physics of the sectioning process.
R. Phillips, Aeon Laboratories, Egham.
47. Interpretation of images of thin sections of keratin fibres.
J. Sikorski, Textile Physics Laboratory, University of Leeds.
48. A survey of embedding media and their physical properties.
A. M. Glauert, Strangeways Laboratory, Cambridge.
49. Molecular behaviour during polymerization.
K. Little, Nuffield Orthopaedic Centre, Oxford.
50. Topochemical studies in relation to the fine structure of keratin.
M. G. Dobb and J. Sikorski, Textile Physics Laboratory, University of Leeds.
51. Rearrangements of dislocations during electropolishing.
U. Valdre, Instituto di Fisica, Bologna.

Summarized Proceedings of a Conference on Thermoelectricity—Durham, July 1961

Abstract

This report gives a summary of the main features of the papers presented at a Conference held by The Institute of Physics and The Physical Society in the University of Durham on 10th, 11th and 12th July 1961.

THERE is a great current interest in thermoelectricity, stemming from the potential application of suitable materials to thermoelectric energy conversion and to thermoelectric refrigeration. For European workers the Durham conference was very well timed coming as it did half-way between the two International Semiconductor Conferences, Prague 1960 and Exeter 1962. The organizers had intended holding a small specialist conference of some eighty participants, but there were so many applications that this number had to be doubled. Thirty came from European countries outside the U.K. and seven came from the United States. It will, regrettably, not be possible in a short report of this kind to attempt to do justice to the thirty-six papers that were read, but it is hoped that many of them will be published independently.

The conference opened on Monday morning, 10th July, most of the participants having arrived the previous day and settled into their accommodation in Grey College and University College (Durham Castle). Professor D. A. Wright (Professor of Applied Physics in the Durham Colleges) who was the chief organizer of the conference welcomed the delegates and opened the meeting with a timely survey of the present state of the subject. From the point of view of applications the work divides into two main subsections: refrigeration, in which alloys based on Bi_2Te_3 still lead the field, and thermoelectric generation, where suitable materials are still being sought. PbTe is the most established material although GeTe and CdSb deserve mention. The big need here is for a material with a large enough energy gap to be extrinsic at high temperatures—for 1000° C it must be over 0·6 ev—as well as having a high z. z is the figure of merit used throughout the conference defined as $z = \alpha^2\sigma/k$ which is proportional to $\mu m^{*3/2}T^{3/2}/k_L$ for a broad-band semiconductor. Here α is the Seebeck coefficient, σ the electrical conductivity, k the thermal conductivity and k_L its component due to lattice vibrations, μ is the mobility and m^* the effective mass of the charge carriers, and T is the absolute temperature. The Professor concluded by outlining some aspects of the general problem. To obtain a high z (or a high zT product) contradictory properties are required in a material. In the first place μ falls as T rises in all broad-band semiconductors. Secondly, although μ rises and k_L falls with increasing atomic weight of the constituent atoms of a compound, the binding energy falls at the same time and thus the energy gap of the materials most suited to thermoelectric application is generally small. The hope of finding a really exciting thermoelectric material seems to be diminishing. Gd_3Se_4 has recently been hailed as a solution, but no one at the conference supported this material.

Dr. R. R. Heikes (Westinghouse, U.S.A.) divided thermoelectric materials into four broad areas: (1) metals, (2) broad-band semiconductors (band width large compared with kT), (3) narrow band semiconductors (band width of the order of kT), and (4) mixed valency semiconductors. A graph of the thermoelectric efficiencies with the years showed that the rate of progress is slowing down. Efficiencies of 10% achieved by 1957 have only reached 18–19% today and the improvement since 1960 has been marginal. He stated that a graph of $\ln \mu$ against $\ln m^*$ showed that for most materials $\mu \propto (m^*)^{-3/2}$ very roughly, thus making z more or less independent of μ and m^* on balance. Several materials exist with a zT of around unity, but none yet exceed 1·5. He ventured the opinion that cubic materials held out the best hope because due to the higher symmetry of the band structure of these materials multivalley semiconduction often occurs and this improves the figure of merit. The occurrence of elliptical energy surfaces can also be shown to be advantageous and this in his view was more important than high atomic weight.

Dr. Ziman (Cambridge) discussed the theoretical interpretation of lattice thermal conduction and explained Umklapp processes. He discussed the influence of isotopes on the thermal conductivity of elements and went on to draw conclusions about the possible effects of impurities in the lattice. He showed that the effect of solid solution is not as advantageous as had been first thought since the lattice thermal resistance varies as $[c(1-c)]^{1/2}$ and not as $c(1-c)$. This has been shown conclusively for LiF with varying Li content.

A great deal of detailed work on individual compounds and alloys was reported and only an outline of this can be presented here.

Dr. H. J. Goldsmid and his co-workers from the Hirst Research Centre of the General Electric Company reported recent work on the anisotropy of electrical properties in Bi_2Te_3. Using a four probe method, they showed how to eliminate errors due to cracked specimens which occur frequently because of the easy cleavage of Bi_2Te_3. Dr. Hérinckx and his co-workers in Brussels described an apparatus they have designed which enables crystals of alloys to be pulled from a melt which has a different composition from the crystal and the bulk liquid feed. The pulling bath is connected to the bulk bath by a fine capillary tube. This apparatus has been used successfully with Bi_2Te_3–Sb_2Te_3 alloys to yield single crystals. Dr. Haacke (A.E.G., Frankfurt) reported work on ageing in copper-doped Bi_2Te_3. Changes in sign of the thermoelectric power are related to the deposition of copper during ageing.

Dr. T. C. Harman (Massachusetts Institute of Technology, Lexington, U.S.A.) reported that a z of $1{\cdot}3 \times 10^{-3}$ has been achieved in HgTe–HgSe alloys near the mid point where k_L is 7 mw cm^{-1} deg^{-1}. Measurements of σ and R have been made from 4·2° K to 350° K and of K and α from 100° K to 300° K.

Dr. J. Shields (Associated Electrical Industries, Rugby) reported work on PbTe, in particular showing slides of slip bands produced by diamond indentation. These compare well with work on NaCl which has the same structure and

indicate that there is some ionic character in the binding of PbTe. An empirical law was established for thermoelectric power and resistivity: $\alpha = 200(4 + \log \rho)$. μ_h was found to be about 1 000 cm^2/v with 10^{18}–10^{19} carriers. In the exhibition which was arranged in the Applied Physics Department during the Conference, Dr. Shields and his colleagues demonstrated a 10-watt thermoelectric generator using PbTe as the n-type junction material, and GeTe as the p-type.

Dr. J. D. Wasscher (Philips, Eindhoven) showed that for a well-behaved semiconductor with μ, r and Ni independent of n and mean free path $l = f(T)E^r$, it is possible to define a constant $\beta = (z_{max}T)_{optimum}$ which may be determined from K, σ and α for a single specimen. Thus it is necessary to make only these three measurements to establish the optimum value of z for a material. He reported that β^* for SnS was 0·30 and that it fell in alloys with SnSe, although the thermal conductivity of Sn_2SSe is 12 mw cm^{-1} deg^{-1} compared with 17 for SnS. SnS is anisotropic and μ, k and z are all higher at right angles to the c axis than along it.

Dr. L. Stourac of Prague was welcomed as the only member of the conference from Eastern Europe. He reported work on CdSb doped with Ag, In, Ga (n-type) and Sn, Pb, Se, Te (p-type) to 10^{-1} at.%. The value obtained for m_p from α and R measurements is rather higher than the accepted cyclotron resonance values.

Dr. B. R. Marathe (Associated Electrical Industries, Aldermaston) contributed at very short notice a paper on $CdSnAs_2$ and $ZnSnAs_2$ and Dr. T. C. Harman and Dr. E. Steigmeier (E.T.H., Zurich) both said a few words about current research at their respective laboratories. R. M. Ware and D. J. McNeil (Plessey, Towcester) reported work on transition metal silicides; these materials have a high conductivity and the work was interpreted on the basis of a metallic-type band structure. Work on $AgSbTe_2$ and the Ag–Sb–Te system was reported by Dr. G. Offergeld and his co-workers (Brussels), by Dr. A. Stegherr (Philips, Aachen), and by Mme H. Rodot (Centre Nationale de la Recherche Scientifique, France). They showed the phase diagram difficulties involved in studying three element compounds.

Two element compounds can sometimes be complicated enough as the session on In_2Te_3 and related compounds showed. Dr. B. R. Pamplin (Durham), Dr. D. R. Mason (Michigan, U.S.A.) and Mr. P. C. Newman (Mullards, Surrey) all gave papers on aspects of the phase diagram, structure and properties of In_2Te_3 and its analogues. Dr. J. E. Parrott (Associated Electrical Industries, Aldermaston) reported theoretical and practical work on the InAs–In_2Te_3 system and Dr. J. C. Woolley (University of Nottingham) reported work on the other alloys of III–V–III_2VI_3 compounds. This session proved to be the fullest session of those devoted to specific materials. Interest in these materials stems from the low thermal conductivity of the defect structure compound In_2Te_3 compared with compounds of its neighbours in the periodic table such as InSb and CdTe. Unfortunately its electron mobility is very low and the increase which occurs on ordering the vacancies is not sufficient to make it of value as a thermoelectric material. Alloys of In_2Te_3 with the higher mobility members of the III–V and the II–VI compounds are, however, worth fuller investigation. Of theoretical interest is the problem of the structure of In_2Te_3 and the phase diagram of the In–Te system in the region near this compound. Both of these problems have now been studied in some detail, but further work is required.

In the session on methods of measuring semiconductor parameters Dr. J. Schröder (Philips, Aachen) described his interesting method of determining the thermal conductivity of a solid. The sample is placed between two boiling liquids so that the temperature difference across the specimen can be known with high accuracy. Measurement of a sample takes only about 10 minutes and with suitable precautions with regard to heat losses the thermal conductivity can be found with an accuracy to ±3%. The method is specially suited to the measurement of a large series of specimens such as is needed to study an alloy system. Miss A. D. Stuckes (Associated Electrical Industries, Manchester) compared the various possible methods of measuring the thermal conductivity of semiconductors at high temperatures: the absolute or axial flow method, the Ångström method and the comparative method in which the axial flow method is adapted so that the unknown material may be compared with two samples of known conductivity. She gave reasons for preferring the comparative method—its main drawback is that a suitable standard must be available. However, when this problem is overcome subsequent results are much more reliable than those of the other methods. L. E. J. Cowles (Hirst Research Centre) read a paper describing apparatus based on the Harman method for measuring the three parameters involved in the figure of merit z. Mr. J. I. Bramman (Atomic Energy Research Establishment, Dounreay) discussed a method of measuring thermal conductivity which was quite novel. If a sphere is placed in a pile where there is a strong γ-ray flux a temperature difference is established between its centre and the outer surfaces. This can be measured quite accurately and admits a very simple determination of K provided the γ-ray flux density is known.

On applications, Dr. K. H. Spring (Central Electricity Generating Board) read a paper describing the three ways in which thermoelectric generators may perhaps be used in large-scale power generation. The chances of replacing conventional plant seem slight though it is possible that in-pile thermoelectric generation might be possible as the technology of this idea develops. This leaves 'tailing' and 'topping', in both of which the output of the thermoelectric stage would be added to the conventionally produced power. Tailing would not be economical because of the small temperature difference involved, but topping is still a possibility for the future provided a suitable material for working near 1000° C can be found. The capital available for such a topping device could be 50% of the plant if a zT of unity at 1400° C could be achieved. Dr. Spring was, however, pessimistic about the future of thermoelectric generation.

All the remaining device papers discussed the theory and practice of thermoelectric refrigeration and in the exhibition in conjunction with the conference several such devices from the General Electric Company were on show. Dr. J. E. Parrott and Dr. A. W. Penn (Associated Electrical Industries, Aldermaston) read a paper on the design theory of a single stage thermoelectric cooling unit and Dr. U. Birkholz (A.E.G., Frankfurt) showed diagrams for optimizing the performance of such units. Dr. W. H. Clingman reviewed research on thermoelectric device design and construction at Texas Instruments Incorporated. Mr. A. R. Sheard (Semiconductor Thermoelements Ltd.) compared zone melted and sintered Bi_2Te_3 and was strongly in favour of the sintered product because with the recent advances in powder metallurgy this seems the most economical method of making thermoelements with good mechanical properties, and the powder product is not found to be inferior in thermoelectric performance to its single crystal counterpart. Dr. J. E. Thompson (English Electric Company Ltd., Stafford) discussed methods of making contact with the cold chamber.

Thermal and electrical contact must be optimized for efficient refrigeration. The junction resistance must not exceed 5×10^{-5} ohm/cm^2. He concluded by emphasizing the need for a cheap thermoelectric material.

The conference was not without its lighter side. Many conference members explored the ancient city of Durham, looking over the Cathedral and the Castle and walking along the banks of the beautiful curve of the Wear on which Durham was founded centuries ago. A coffee evening held in the Castle was well attended and Dr. Prowse, vice-master of University College and himself a physicist, showed members round the Castle with its Norman chapel and many interesting architectural features and historic treasures.

In spite of a note of pessimism for the prospects of thermoelectric application on a large scale, the conference was a happy and successful one and all appeared to enjoy their visit to Durham.

Department of Applied Physics,
Science Laboratories,
The Durham Colleges,
South Road, Durham City.

B. R. PAMPLIN

The following papers presented at the Thermoelectricity Conference have been accepted for publication:

Measurement of the figure of merit of a thermoelectric material.

Miss A. E. Bowley, L. E. J. Cowles, G. J. Williams and H. J. Goldsmid (*Journal of Scientific Instruments*, **38**, 433).

Some problems in the development of a commercial thermoelectric refrigerator.

T. B. Burnett, H. O. Lorch, and J. E. Thompson (*British Journal of Applied Physics*, **12**, 595).

Anisotropy of the electrical conductivity in bismuth telluride.

R. T. Delves, Miss A. E. Bowley, D. W. Hazelden and H. J. Goldsmid (*Proceedings of the Physical Society*, **78**, (*5b*)).

The measurement of thermal conductivity of semiconductors at high temperatures.

Miss A. D. Stuckes (*British Journal of Applied Physics*).

Some problems in the development of a commercial thermoelectric refrigerator

by T. B. BURNETT, M.Sc., A.M.I.E.E., H. O. LORCH, A.M.I.E.E., and J. E. THOMPSON, B.Sc., Ph.D.,
The English Electric Co. Ltd., Nelson Research Laboratories, Stafford

MS. received 22nd June 1961; paper presented at the Conference on Thermoelectricity held by The Institute of Physics and The Physical Society, at Durham, 10th–12th July 1961

Abstract

Two basic problems in the development of a thermoelectric refrigerator are the construction of a cooling unit, consisting of the thermoelectric junctions, and the optimization of the complete refrigerator design to make best use of the thermoelectric material. It is essential for the unit to achieve good thermal contact with the cooling chamber and to have good electrical contact across the p and n semiconductor bars forming each junction. A simple but accurate method for measuring the junction resistance is shown, and the relative merits of various methods of forming the cooling unit are discussed. The heat balance equations for a refrigerator are given which include the cooling fin system, and the method of optimization of the design is noted.

1. Introduction

PELTIER cooling as a means of refrigeration for specialized devices is well developed, and for commercial refrigeration is worthy of active study. Present materials are marginal in value for most devices, except those requiring low heat extraction rates; but there are obvious advantages linked to a Peltier cooling system. A study of thermoelectric cooling problems, the type of material and the production and design of cooling units is pertinent at the present time.

Bismuth telluride alloys are at present the most commonly used materials in thermoelectric cooling units. These materials can be made in a number of ways. To achieve optimum thermoelectric properties, the material is zone refined in either the element or compound form, then doped to the optimum impurity concentration and finally zone levelled to produce uniformity of properties along the ingot length. During the process some tellurium is lost from the melt, but allowance can be made for this by a trial and error method. Such material is anisotropic in properties, large grained, and easily broken along the cleavage planes. Individual crystals have the c axis perpendicular to the direction of zone refining (Goldsmid 1958). In one method used in the construction of a cooling unit, the ingot is cut into bars of suitable cross-sectional area and length, and finally p and n bars are joined together in a jig by metal strips to produce a cooling array. It is essential that low resistance joints are made between bars so that no thermal loading of the cold junction takes place, and that the final unit has flat end faces. If the bars are accurate in length this can be achieved easily, but if not it is necessary to grind the end faces of the final unit.

Alternative methods to be described of constructing a cooling unit have sought to improve the mechanical strength of the bars by using small grained material—as for instance by sintering—to produce low junction resistances across the ends of the bars by improved joining techniques, and by a technique of r.f. casting to produce bars of the correct length, small grain size, and with cast-in end junction pieces. For the measurement of this junction resistance a simple method is described, which also attempts to ensure uniform current distribution in the bar.

In any adequate design for a commercial thermoelectric refrigerator, to give best utilization of the properties and quantity of material employed, the heat balance equations must be considered for the whole system. The derivation and method of solution of such equations are discussed.

2. Construction of the cooling unit

The construction of a satisfactory cooling unit requires that the thermoelectric bars be mechanically strong, that joints of low resistance can be made to the bars, and that good thermal contact can be achieved between the unit and the refrigerator walls and fin system.

A number of techniques will be described which, with varying degrees of success, overcome these problems.

2.1. *Material produced by cooling the melt in small bore tubes*

One method of obtaining strong bars, which eliminates a great deal of cutting, is to melt powdered bismuth telluride contained in small bore tubes by lowering the tubes through a short hot zone in a furnace. Quartz tubes ($\sim \frac{1}{4}$ in. dia.) can be used, and bars of accurate length are obtained by slitting the tube lengthways and then cutting the rod into pieces of the correct length. There is a limit to the diameter of the bar since the crystal size increases with increasing diameter of the tube. Above some size dendritic growth and impurity segregation occur, both producing non-uniform properties. At misaligned grain boundaries and regions of segregation excessive heating or cooling can take place.

The importance of this segregation in various techniques has been discussed by Ainsworth (1956), who studied the growth of single crystals by various methods. Satterthwaite and Ure (1957) have obtained crystals of almost uniform composition (using a specially shaped container) by crystallizing very slowly from a large melt, so that the excess of one component rejected at the crystallizing surface diffuses throughout the melt.

However, with the method used here, where a sealed tube of bismuth telluride of the correct properties is melted and cooled rapidly through a hot zone of a furnace, segregation is reduced to a minimum and small grain size is obtained.

Thermoelectric cooling units are often made from this

form of material by taking these small cut circular bars, 'trimming' the ends, holding them in a suitable array in a jig (with metal connecting end pieces over the junctions), and heating the assembly under slight pressure to a temperature above the melting point of a solder. With care and the correct amount of solder a firm array of cooling junctions can be formed. For a solder Joffe (1957) suggests an eutectic of 80% bismuth, 20% tin with a melting point of 200° C, and a German patent (858925, 21st July 1950) gives a solder consisting of 51% indium and 49% zinc which melts at 117° C for use with low melting point alloys. Of the various solders tried the one found to be the best consisted of bismuth, tin, antimony and silver alloy, which had a liquidus temperature of 271° C and a solidus temperature of 185° C. With a suitable flux 'Rosinal' this would wet bismuth telluride with or without an ultrasonic soldering iron.

To produce perfectly flat smooth end faces, the faces are usually ground, and to provide electrical insulation (but good heat transfer) a thin layer of insulating material is applied between the junctions and the final thin metal end plates. For this Joffe recommends silicone varnish with 6% aluminium powder, though a thin sheet of natural rubber is quite satisfactory.

2.2. *The use of sintered material*

In sintering the elemental powders, or more usually the alloy in finely crushed form, are compacted together and heat treated below the melting point. Diffusion takes place between particles and a small grained, almost homogeneous bar results. Shrinkage takes place during sintering, but, with care, the amount is reasonably constant and the small grain size gives increased strength. The sintering method has been described by Vasenin (1955) for various alloys of antimony and tellurium.

Although soldering to sintered bars is normally quite difficult, the situation is improved considerably by the inclusion at the ends of the bar of a suitable metal facing which is wetted by the solder. The metal powder is introduced at the ends of the bar, either during the initial pressing, or after the sintering process has been carried out. In the latter case, re-pressing of the block with metal powder at the ends and a subsequent low temperature heat treatment are necessary to give a good bonding of the bar and metal.

Silver is an obvious choice for a metal addition. However, when the experiment was tried the resulting surface could not be wetted and the silver had apparently disappeared. On examination, white needle-like crystals were found at the compact ends (Fig. 1), which were identified as silver telluride by x-ray diffraction. From the phase diagrams of silver–bismuth and silver–tellurium, bismuth silver eutectic forms at 262° C and a silver–tellurium eutectic at 351° C. Silver and tellurium do not react at a measurable rate in the solid state,

Fig. 1. Photomicrograph of sintered bismuth telluride with metal end contacts. Crystals of silver telluride can be seen as thin needles. Metallographic preparation: wet grinding through papers grades 240–600 and final polishing using 6 micron and 1 micron diamond pastes on nylon and Selvyt cloths respectively. Etchant: cupric chloride 6 g, hydrochloric acid 1 cm³, alcohol 90 cm³.

so that the formation of silver telluride would not be expected below 351° C, but the presence of the bismuth–silver eutectic allows the reaction between silver and tellurium to take place above 262° C. Sintering below 262° C would eliminate this problem, although the silver itself would not bond at this temperature. Zinc proved to be the most successful metal. When bismuth telluride compacts were sintered, and then subsequently heated to bond on the zinc powder, a good key was obtained; on breaking, the bar fractured, but not the joint.

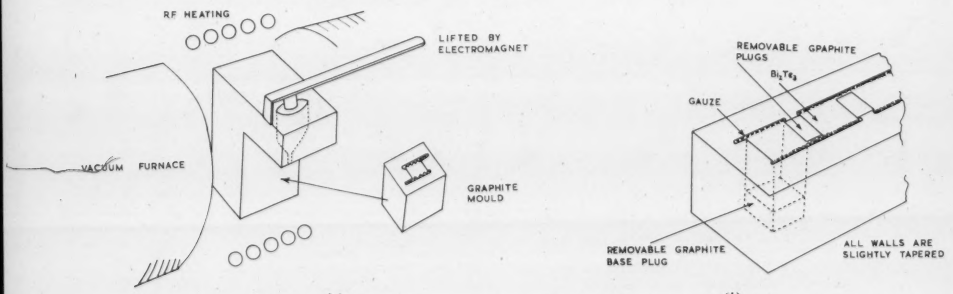


Fig. 2. The casting technique for producing blocks of the correct size with cast-in end connections.
(*a*) Production of single blocks. (*b*) Production of an array.

2.3. *The r.f. casting of an array of elements*

One way of producing small grained blocks of the correct size that has received serious attention is the casting of blocks from the melt. In this method, crushed bismuth telluride is melted by r.f. inducton heating in a graphite container situated in a vacuum chamber (Patent submitted). A graphite plug in the bottom of the furnace can be raised or lowered by a metal arm operated by an electromagnet outside the chamber. The released bismuth telluride then runs into a small mould which gives rapid cooling and consequently small grain size (Fig. 2(*a*)). Apart from the fact that the cast bars produced are of the correct size, and so require no cutting, this method offers other advantages. End connections are made by arranging for metal gauze to be held at the end of the mould, so that, provided a metal that is wetted by bismuth telluride is chosen for the gauze, the molten metal flows through the gauze. Fig. 3 shows a micrograph of such an end piece. Since the bar mechanically embraces the gauze a firm joint is made. A metal has to be chosen which does not diffuse rapidly into the molten bismuth telluride and thereby alter its properties.

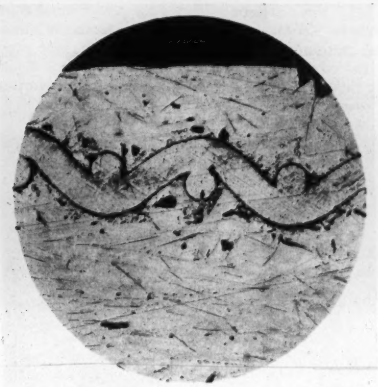


Fig. 3. Photomicrograph of a sectioned block showing the excellent keying of the metal to the end connection.

The greatest advantage of this method lies in its commercial application. Using p and n melts, alternate bars can be cast into alternate holes in an extended mould to produce a series of thermojunctions complete with electrical connections (Fig. 2(*b*)). Moreover, the bars are mechanically strong and are of the correct length. The short molten period eliminates the loss of constituents from the melt and subsequent change of thermoelectric properties. Larger units involve only the soldering together of end connections.

3. The measurement of junction resistance and its importance

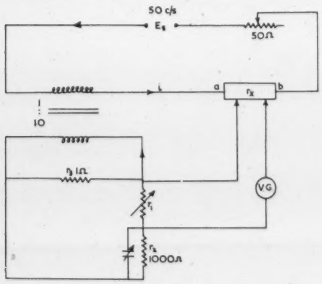
In any method of forming cooling arrays it is necessary to show that the junction has low electrical resistance. This arises since the net heat extraction rate from the cold end is small compared with the rate at which heat flows back to the cold end through the bars themselves. For example, in one particular design with a cooling current of 20 A, a junction resistance of greater than 50×10^{-6} ohm could not be allowed. One method of measuring junction resistance or resistivity is by use of the modified Wheatstone bridge network of Goldsmid (1958) which he used to measure the electrical conductivity of bismuth telluride.

A quick and accurate method of resistance measurement is shown in Fig. 4. A.C. current is passed through the thermoelectric cooling element and a current transformer in this circuit provides a voltage across a standard resistance r_3 (Fig. 4(*a*)). A variable portion of this voltage is fed back to balance the voltage picked up by two potential probes across the specimen a known distance apart. It is simple to show that provided $r_2 \gg r_1$, the resistance across the probes r_x is given by

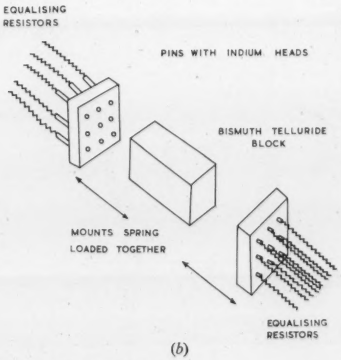
$$r_x = \frac{r_1 r_3}{N r_2}$$

where N is (secondary turns/primary turns) of the current transformer.

When measurements of resistivity were made on these blocks using this circuit, and current contacts consisting of a



(*a*)



(*b*)

Fig. 4. The measurement of resistivity and junction resistance.
(*a*) The basic circuit. (*b*) The method used to give equalization of the current density into the block.
V.G., vibration galvanometer 50 c/s.

thin pad of indium pressed into the ends by copper blocks, widely varying values of resistivity were found in different parts of the block. It was not certain whether these results were due to inhomogeneities in the block or failure of the indium pads to make uniform contact over the ends of the block. To study this further, the single contact pads were replaced by nine contacts at each end, which consisted of small pads of indium affixed to pin heads. The contacts were connected in series with resistances of larger resistance than the bismuth telluride blocks (and contact resistance), so that each contact carried approximately the same current (Fig. 4(*b*)). Measurement of resistivity with this method also showed variations, indicating that the inhomogeneous nature of the blocks was responsible for the variation in resistivity. In a typical case of a junction to zone refined material one pick-up connection was made to the bar end. The other end, mounted on a micrometer attachment, was placed close to the joint in several places in succession to obtain a representative value of resistance.

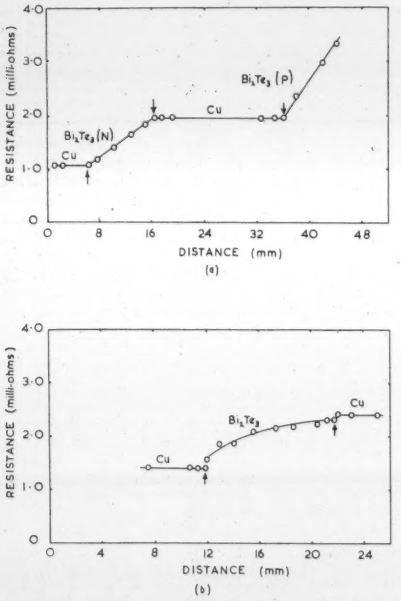


Fig. 5. Resistance probe measurements along a bar showing diffusion and modification of the properties of the bar, indicating (*a*) good joints, and (*b*) poor joints and the effect of impurity diffusion into the bar.

Certain junction materials and solders are able to diffuse into the material during the soldering process and cause deterioration in properties. Resistance probe measurements taken along the bars are shown in Fig. 5. Good joints are illustrated in Fig. 5(*a*), where the arrows indicate the positions of the junctions. Fig. 5(*b*) illustrates the effect of poor joints, and the change of properties of a specimen along its length due to impurity diffusion from the joint during the soldering process.

4. The design of a refrigerator

Although bismuth telluride and its alloys are the best thermoelectric materials known, their properties are only marginally useful except for small specialized applications, but a refrigerator was designed to examine its commercial possibilities.

The maximum net pumping power of the cooling unit is proportional to the total area to current flow, and subdivision of this area into small elements, electrically in series, increases the impedance and hence the efficiency of the power supply. The practical limit is set, not by this impedance, but by the difficulties of making and joining many very thin elements. For domestic purposes the power input is limited to perhaps 100 w, as the refrigerator necessarily must take power continuously while in use. The net heat extracted from the container or 'ice box' is then far less dependent on the total area of the elements because the current through the elements, which form a large total area, is limited by the available power input to a value below the optimum value.

Very efficient cooling of the hot junction is necessary to limit its temperature difference from the air temperature, but aesthetic considerations limit the overall size of the cooling fins. The fin spacing (Fig. 6) is subject to optimization of

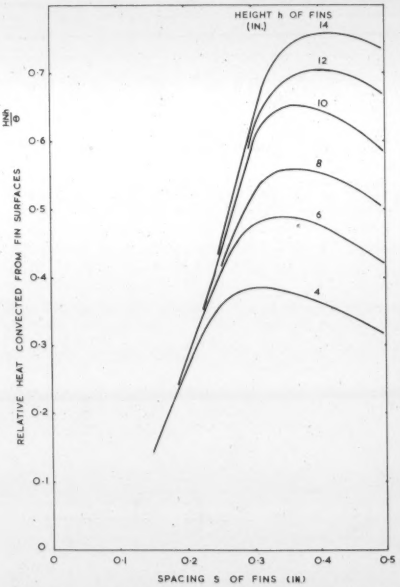


Fig. 6. Showing the influence of fin dimensions on the loss of heat from the fin surfaces.

H in w/in^2, N is the number of fins on a plate 12 in. wide, h is the height of fins (in.), θ is fin to ambient temperature difference (deg. C). The optimum spacing occurs when $s/\{(Gr)(Pr)\}^{0\cdot25}/h = 2\cdot75$, where ($Gr$) and ($Pr$) are Grashof and Prandtl numbers.

design. If the number of vertical fins in a fixed width is steadily increased the total surface area convecting increases nearly in proportion to the number, and at first the total heat convected rises, but beyond a certain optimum spacing the

air currents from one fin surface interfere with those from the next to such an extent that there is less total heat convected, until when the fins are about 0·1 in. apart convection currents almost cease. For optimum spacing the heat convected per deg C temperature difference is proportional to the width and depth, and the square roots of the height and temperature difference.

The amount and quality of the lagging around the ice box depends on the cost of the lagging and overall size of the box, and differs little from conventional practice except that especial care is needed in insulating the warm back plate of the cooling fins around the cooling unit from the ice box.

4.1. *Analysis of operation*

Several authors (e.g. Altenkirch 1911 and Joffe 1957) have given heat balance equations for individual pairs of elements which may be optimized, in terms of voltage or current, for maximum efficiency or heat extraction rate when one end is maintained at constant temperature.

In a commercial refrigerator it is necessary to include many other terms in the heat balance equations, which has been done in a simple form by Lackey *et al.* (1958), and Penrod (1960). Nevertheless, a more comprehensive method of designing a refrigerator is as follows.

The thermal circuit is shown in Fig. 9. The heat W_1 leaking into the ice box through its lagging (R9 and R10) is pumped away by the cooling unit through the cooling fins and back to ambient air. The temperature obtained in the ice box depends on

(i) the net pumping power of the cooling unit;
(ii) the box lagging (R9 and R10);
(iii) the convection and radiation resistance R11 of the fins to air;
(iv) the resistance R7 and R8 between the cooling unit, ice box and fins,

and will be determined by the solution of the equations which follow.

4.2. *Symbols*

The following symbols, expressed in watts, inches, °C, ohms and amps exclusively, are used in the equations:

Temperature T_0 of the ambient air, T_1 of most of the top, bottom sides and front of the ice box, T_2 of most of the back of the ice box, T_3 of the cold junctions, T_4 of the hot junctions, T_5 of the fins.

N number of cold junctions, i.e. half number of elements.
a cross-sectional area of one element.
a_1 area per junction for heat flow between elements.
l length of each element.
I current through each element.
α thermo e.m.f. of a junction pair.
ρ sum of resistivities of a p and n element pair.
k sum of thermal conductivities of a p and n element pair.
w_1 heat extracted from one cold junction, w_2 heat dissipated at one hot junction.
W_1 total watts extracted, W_2 total watts dissipated, W_3 electrical power input to the cooling unit.

Electrical resistance R_1 of the joints at a cold junction, R_2 of the joints at a hot junction. Thermal resistance per unit area R_3 of the cold junction to ice box, R_4 of the hot junction to fins. Thermal resistance R_9 between sides, front, top and bottom of ice box and the ambient air, R_{10} between back of ice box and the fin back plate.

4.3. *The heat equations*

The heat pumped from one cold junction is, by definition, $\alpha(T_3 + 273)I$. Half the ohmic loss in each element is returned to the cold junction and half adds to the hot junction, so the gross ohmic loss at a cold junction is $I^2(\rho l/2a + R_1)$.

The heat conducted back through the two elements is $(T_4 - T_3)ka/l$ and the heat convected and radiated back within the space between the elements, which is small in comparison, can be shown to be roughly

$$(T_4 - T_3)a_1\left(0{\cdot}0014 + \frac{0{\cdot}0008}{l}\right).$$

Hence

$$w_1 = \alpha(T_3 + 273)I - I^2\left(\frac{\rho l}{2a} + R_1\right) - (T_4 - T_3)\left(\frac{ka}{l} + 0{\cdot}0014\,a_1 + \frac{0{\cdot}0008a_1}{l}\right). \quad (1)$$

Similarly

$$w_2 = \alpha(T_4 + 273)I + I^2\left(\frac{\rho l}{2a} + R_2\right) - (T_4 - T_3)\left(\frac{ka}{l} + 0{\cdot}0014a_1 + \frac{0{\cdot}0008a_1}{l}\right). \quad (2)$$

The area of a bar joining the cold ends of two elements is about $2a + 0{\cdot}2a_1$ and so for heat flow from the ice box to the cold junction we have

$$T_2 - T_3 = \frac{w_1 R_3}{2a + 0{\cdot}2a_1}. \quad (3)$$

Similarly for the hot junctions

$$T_4 - T_5 = \frac{w_2 R_4}{2a + 0{\cdot}2a_1}. \quad (4)$$

The heat extracted passes through the walls and top and bottom of the ice box giving

$$Nw_1 = (T_0 - T_1)R_9 + (T_5 - T_2)R_{10}.$$

To simplify the calculations this was approximated to by the simple expression

$$Nw_1 = (T_0 - T_2)\frac{R_9 R_{10}}{R_9 + R_{10}}. \quad (5)$$

The calculated sum of the convection and radiation from the fins was plotted against temperature, and it was seen that one simple expression would suffice to cover the region of interest with sufficient accuracy. This expression, when corrected slightly by a measurement on the actual fins at one temperature, is

$$Nw_2 = 5{\cdot}05(T_5 - T_0) + 0{\cdot}098(T_5 - T_0)^2. \quad (6)$$

The ice box is made of sheet metal, and it is found that T_1 may be expressed approximately by

$$T_1 - T_2 = 0{\cdot}25Nw_1. \quad (7)$$

Three more subsidiary equations are obvious

$$W_1 = Nw_1, \quad (8)$$

$$W_2 = Nw_2, \quad (9)$$

$$W_3 = W_2 - W_1. \quad (10)$$

The solutions of the simultaneous equations (1) to (6) and the subsidiary equations (7) to (10) give all the temperatures and the quantities of heat transferred in the refrigerator, for any chosen current.

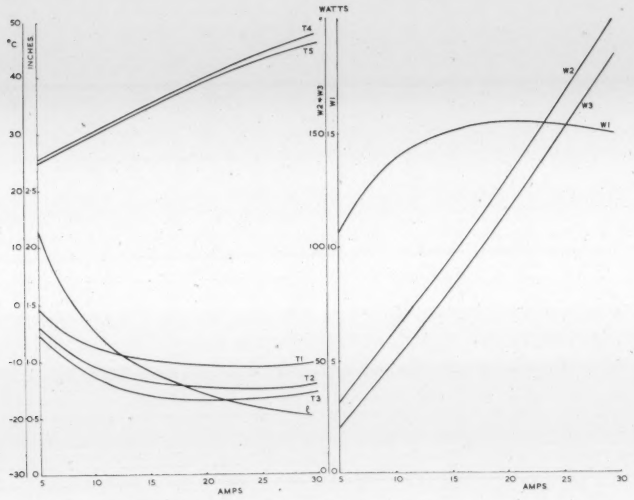


Fig. 7. Graph showing the dependence of various parameters on the current flowing.

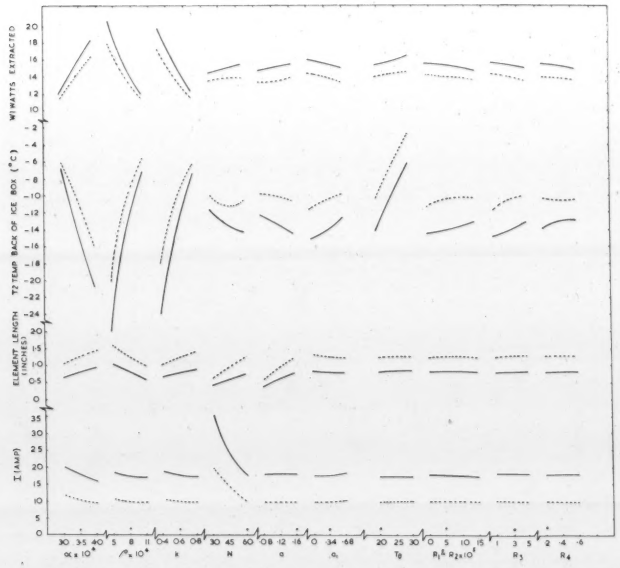


Fig. 8. The effect on performance of varying 10 parameters in turn from a nominal value and keeping the remainder constant. ——— 100 w input – – – – 50 w input. Nominal values are marked by a dot.

The optimum length of element to give the greatest heat extraction occurs when the ohmic loss returned to the cold junction is equal to the heat conducted back, and is independent of the joint resistances. It is given by

$$l = \frac{a}{I}\left\{\frac{2k(T_4 - T_3)}{\rho}\right\}^{1/2} \qquad (11)$$

if the small air terms are ignored. Putting (11) into (1) and (2) gives

$$w_1 = I\left[(T_3 + 273)\alpha - \left(2 + \frac{0{\cdot}0008a_1}{ka}\right)\left\{\frac{\rho k(T_4 - T_3)}{2}\right\}^{1/2} - IR_1\right] - (T_4 - T_3)0{\cdot}0014a_1 \qquad (12)$$

$$w_2 = I\left[(T_4 + 273)\alpha - \frac{0{\cdot}0008a_1}{ka}\left\{\frac{\rho k(T_4 - T_3)}{2}\right\}^{1/2} + IR_2\right] - (T_4 - T_3)0{\cdot}0014a_1. \qquad (13)$$

The importance of each of the parameters T_0, N, a, α, ρ, a_1, R_1, (R_2 and R_3) and R_4 was studied by solving equations (3) to (13) (i.e. with optimized length) for 5, 10, 20 and 30 A on a DEUCE computer. It was not practicable to solve for all possible combinations, so all parameters were given a 'nominal' value and one at a time varied to two other values, keeping the rest nominal.

Figure 7 shows the solution for all parameters nominal, and a similar set of curves was obtained from all the 21 sets of solutions. It is seen that there is a broad optimum of current around 22 A, but that the electrical input is then 127 W. This may, however, be reduced to, say, 50 W without much loss of performance.

The four solutions of most interest, w_1, T_2, l and I extracted from the curves at 100 W and at 50 W are presented in Fig. 8 and the conclusions derivable are:

(i) α, ρ, and k are all critically important, as expected.

(ii) The number of junctions may be reduced from 60 to less than 30 without seriously affecting performance. The area of the elements is also surprisingly unimportant and could be reduced to half nominal or less. For these reasons, the quantity of thermoelectric material is drastically reduced in subsequent designs.

(iii) The space between elements is not very important, and neither, within the range considered, are R_1 to R_4.

(iv) The performance relative to room temperature is little affected by the room temperature.

(v) The efficiency at 50 W input is about 28%.

4.4. *The complete thermal design*

The complete thermal design of a commercial refrigerator is shown by the values in Fig. 9, which shows that with 50 W electrical input 14 W are extracted from the ice box to give a temperature drop of about 29 deg C below the room temperature (28% efficiency). In this particular design, the size and number of the elements used was fixed at 120 elements, 0·4 in. square × 0·45 in. long for manufacturing considerations, though later improved designs used a much smaller volume of material—a smaller number of elements of both smaller cross-sectional area and length.

A conventional 'absorption' type of refrigerator of about the same size was found experimentally to extract about 13 W for an electrical input of 104 W (12½% efficiency), but had the advantage of maintaining good thermal insulation when switched off.

To operate the thermoelectric refrigerator a 10 A, 5 V supply is required.

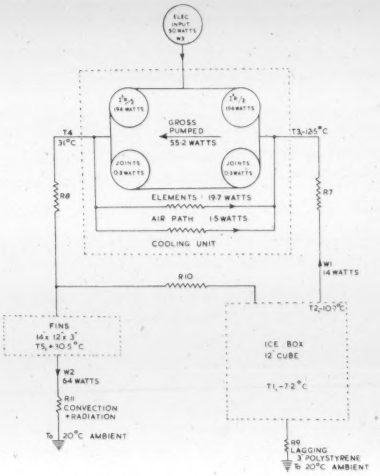


Fig. 9. The complete thermal circuit of a domestic refrigerator.

5. The power supply

In any commercial refrigerator the overall efficiency must be considered; this will include the efficiency of conversion from a.c. mains supply to the d.c. current to operate the thermoelectric cooling element.

A typical full-wave rectifier circuit is shown in Fig. 10, to provide a current in the range 10–20 A at a few volts. The bi-phase rectifier is used since the rectifier voltage presents no

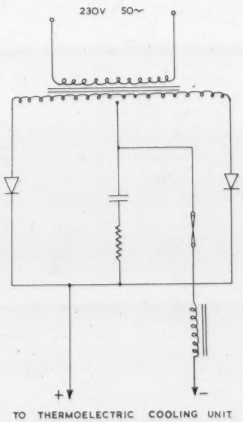


Fig. 10. A typical bi-phase rectifier power supply.

problem, and has the advantage of requiring the minimum number of rectifiers, and the load current flows through only one rectifier arm at a time. As the rectifier forward drop is approximately a constant value at the rated current, the rectifier loss (which is the greatest loss in the converter) depends on the current used. Using the most suitable rectifier, a germanium p–n junction, the voltage drop at rated current is about 0·7 v, so that a power loss of 7 w can be expected at 10 A.

Consequently it would be desirable to operate the cooling unit at as low a current as possible. This however may mean reducing the cross-sectional area of the elements and increasing the number of junctions so that the saving in loss in the a.c. to d.c. converter must be balanced against the increase in manufacturing complexity. In the operation of the cooling unit the heat extracted depends on the average current, while the I^2R loss depends on the r.m.s. current. Any increase in the ratio r.m.s. current/average current—the form factor of the current—above unity works to the disadvantage of the cooling element. As the normal voltage output from the power unit is a rectified sine wave, the form factor of which will be in excess of 1·11, smoothing is necessary to provide a current of an adequately low form factor.

For a form factor of 1·01, which would give a 2% increase in I^2R loss, a ratio of $I_{ac}/I_{dc} = 0{\cdot}2$ is required.

At the low voltages required for supplying the cooling element capacitive smoothing is not practical and so a series inductance has to be used.

The effect of a ripple current has been discussed in some detail by Alfonso and Milnes (1960).

Losses in the smoothing choke and transformer will add about another 10% to the watts to be supplied. Thus a converter to supply 10 A at 5 v to a thermoelectric cooling element would require an input of $50 + 7 + 6 = 63$ w and the overall efficiency is reduced to $14/63 \times 100 = 22\%$. The rectifiers in the converter have to be protected against voltage surges by the R–C circuit shown in Fig. 10.

6. Conclusions

The technical problems in the economic fabrication of cooling elements have been overcome, and the design techniques for a commercial thermoelectric refrigerator are satisfactory. The chief factor limiting the general application of thermoelectric cooling is the cost of producing the basic material for the elements, since bismuth telluride alloys of reasonably high purity and control are required to obtain the best thermoelectric properties. The properties of bismuth telluride are really marginal in value and a material with a higher figure of merit, and with which some control over the purity could be relaxed, is highly desirable.

Acknowledgments

The authors are indebted to R. Sharples, who co-operated in the early stages of the work, and to J. K. Brown, Director of Research, for permission to publish this paper.

References

AINSWORTH, L., 1956, *Proc. Phys. Soc.* B, **69**, 606.

ALFONSO, N., and MILNES, A. G., 1960, *A.I.E.E. Conference Paper No.* 60–174.

ALTENKIRCH, E., 1911, *Phys. Z.*, **12**, 920.

GOLDSMID, H. J., 1958, *Proc. Phys. Soc.*, **71**, 663.

JOFFE, A. F., 1957, *Semiconductor Thermoelements and Thermoelectric Cooling*. (London: Infosearch Ltd.)

LACKEY, R. S., MEES, J. D., and SOMERS, E. Y., 1958, *Refrig. Engr*, **66**, 31.

PENROD, E. B., 1960, *World Refrigeration*, **11**, No. 8, 32.

SATTERTHWAITE, C. B., and URE, R. W., 1957, *Phys. Rev.*, **108**, 1164.

VASENIN, F. I., 1955, *Zh. Tech. Fiz.*, **25**, 1190.

Clean glass substrates for Permalloy films

by A. F. FRAY and S. NIELSEN, B.Sc., Ph.D., Ministry of Aviation, Royal Radar Establishment, Great Malvern, Worcs.

MS. received 30*th November* 1960, *in revised form* 13*th June* 1961

Abstract

Possible defects on glass surfaces and methods of detecting and reducing them are discussed. A correlation between the state of the surface and the magnetic properties of Permalloy films is observed. Best results were obtained on surfaces obtained by quenching molten vacuum-melted glass.

Introduction

Increasing attention is being devoted to the preparation and properties of thin films, and it is widely appreciated that the properties of such films may be markedly influenced by the nature and surface of the substrate. This is certainly the case for evaporated Permalloy films. Glass, being cheap, amorphous and stable at reasonable temperatures, is a very suitable substrate for Permalloy films; microscope slides or cover slips are commonly used. However, the surface of such substrates may possess various types of defect and it is the purpose of this paper to discuss these and to report experiments designed to improve the surface, having special reference to the magnetic properties of Permalloy films.

Surface defects

Defects on glass surfaces may be caused by: (*a*) polishing marks and scratches, (*b*) 'Griffiths' cracks, (*c*) surface dust and grease, (*d*) the gel nature of the surface layer, (*e*) gas and vapour evolution during bake-out and (*f*) contamination in the vacuum system.

Polishing marks, scratches and pits can often be recognized by a direct microscopic examination of the surface, and can be revealed more clearly by the etching procedures. Such marks are presumably introduced during manufacture and during any processing treatment prior to use. In this connection it is suggested that slides should be cut in the presence of water; the electron microphotographs of Joos (1957) show how the trace of scratch changes from a broad track with deep cracks at right angles, to a narrow track with well-defined edges as the presence of water vapour is increased. Certainly, a too heavy pressure on the diamond produces easily visible cracks at right angles to the scratch.

The existence of minute cracks or flaws in glass was postulated by Griffiths in 1920, and provided an explanation for the failure of glass to exhibit its theoretical strength. Recently Levengood (1959) has shown that such flaws can be revealed by etching and that the etch pit pattern is influenced by the internal stresses to which the glass has been subjected, this resembling in many ways the dislocation pattern on crystal surfaces. With sheet glass, for example, the flaws in general lie parallel to the direction of the 'draw'.

Attempts have been made to remove the dust, grease and other chemicals adhering to the surface, by a variety of cleaning procedures. Dust is a possible cause of pinholes in thin films and should be removed from the vacuum system prior to evaporation, as the turbulence during the initial pumping may be sufficient to distribute it over the substrate surface. The removal of grease and chemicals is complicated by the nature of the glass surface, which is thought to consist of a layer of silica gel probably 0·5 μ thick (Joos 1957, Koranyi 1959). This layer may be expected to absorb the vapours and ions to which it has been exposed, and to retain and exchange them in a similar way to a chromatographic column of the same material. Chemical reagents used for cleaning may themselves be absorbed by the gel layer and subsequently be difficult to remove. Chemical reagents may also interact with those cations able to diffuse through the gel layer. Thus Tsuchikaski *et al.* (1958) concluded from their work on heavy crown glass that Ba^{++} ions migrate easily through the gel layer to form a surface layer of $Ba(OH)_2 8H_2O$, $BaCO_3$, $BaCl_2$ and $BaSO_4$, with water vapour, wet CO_2 gas, wet HCl, and wet SO_2 respectively. Any cleaning procedure therefore should take into account the interaction of the cleaning reagent with the 'gel' layer and with the cationic components of the glass. Cleaning reagents should also not attack the surface irregularly, e.g. by preferential attack at 'Griffiths' cracks.

Glass may also contain appreciable quantities of 'water' and dissolved gases. For example, Kruszewski (1959) reports that a 10 g sample of glass contained 1·8 cm^3 of a gas, whose composition was H_2O 74·8, CO_2 6·1, O_2 6·9, CO 3·9, H_2 7·8, N_2 0·3%. Todd (1955) shows that the gas evolved during a vacuum bake-out is mainly water, and that the amount of water evolved at constant temperatures above 300° C is linear with respect to the square root of the bake-out time. He also distinguishes between easily removable surface water and water diffusing from the interior. The amount of surface water depends on previous history. A two-year old glass, which had been previously washed with water, gave about 64 micron litres per square decimetre of surface, compared with about 9 $\mu l/dm^2$ from the same glass washed with 1% HF; (3 $\mu l/dm^2$ is comparable with the water contained in a unimolecular absorbed layer). The amount diffusing from the interior is of course, temperature dependent, about 20 $\mu l/dm^2$ being evolved after four hours bake-out at 480° C.

Also included as a source of defects, is the possible, indeed probable, contamination of the surface in the vacuum system, especially by oil vapours (Holland and Bateman 1960). This is not discussed here, but presumably contamination is at a minimum if the substrate is kept at as high a temperature as possible.

Surface tests

Attempts to estimate the state of the surface have been made in many ways. For example, Holland (1958) and Putner (1959) have used the coefficient of friction, optical

absorption, wettability, breath tests and adhesion of thin films, and in their study of surfaces, Tsuchikaski *et al.* (1958), mention the use of phase microscopy, measurement of haze value, electron diffraction observation of dew drops and electron microscopy. Konig *et al.* (1955) condensed cadmium and titanium dioxide on the glass to show the surface structure, and further showed that the chemical nature of the surface was an important factor in the condensation of metals. Heterogeneous catalysis of gaseous reactions has also provided information on the chemical nature of the surface.

In our experiments we have firstly measured the weight required to produce a visible titanium scratch mark, Holland (1958) having noted that titanium leaves a scratch mark more easily on clean surfaces. The apparatus is illustrated in Fig. 1. Here the glass slide is moved horizontally under the

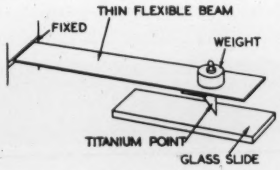


Fig. 1. Apparatus required to produce a visible titanium scratch mark.

titanium point, and the weight above the point increased until a visible mark is produced. This technique gave reproducible results and was surprisingly sensitive to surface contamination.

saturated with NaF was used for five minutes at room temperature.

Finally some of the 'best' slides were examined by an electron microscope replica technique carried out by the Aeon Laboratories, Egham, Surrey. In this examination the surface was shadowed with palladium–carbon and the replica stripped with NaOH.

The results of these tests are discussed after outlining the cleaning procedures used. Subsequently the results are correlated with observed magnetic properties.

Cleaning procedures

Several cleaning procedures were used on 3 in. × 1 in. sections of microscope slides. The first stage in the cleaning attempted to remove surface dust and grease, and the second stage to remove traces of reagents used in the first stage. For the first stage we favoured detergent (Teepol) and water, rubbing with a soft pad or slurry of fine powder, followed by a vigorous wash with hot and cold water. This resulted in a more even wetting of the surface, but the contact angle did not appear to be especially low. The procedure recommended by Yarwood (1956) using chromic acid showed no improvement on this (see the Table), but the action of various acids produced a more evenly wetted surface. However, in view of the possible attack on the glass surface, acids and alkaline were not thought desirable.

Removal of water was attempted by draining and evaporation, blowing with a jet of clean cold air (a technique suggested by workers at the British Scientific Instrument Research Association), polishing with dry cloths and by suspending in ethanol vapour. Removal of water by washing with alcohol, followed by these treatments, was also attempted.

Cleaning procedure	*Weight on Ti(g)*	H_c(*Oe*)	H_k (*Oe*)	*Zinc layer*
(1) None	~400	>8	>8	Much structure
(2) 10 min in trichloreth, acetone, H_2O, 10% NaOH, conc. HNO_3, H_2O, drain dried	100	6–7	~8	Much structure
(3) Teepol swab, H + C H_2O, conc. HNO_3, ethanol, flame dried	50	5–6	7–8	Structure where alcohol evaporates
(4) Teepol swab, H + C H_2O, chromic acid (as Yarwood 1956)	100			Much structure
(5) Teepol swab, H + C H_2O, drain dried	100	5–7	7–8	Much structure
(6) Teepol swab, H + C H_2O, ethanol, drain dried	100	5–7	7–8	Much structure
(7) As (6) + wiped with clean Selvyt cloth	60–100	3–4	4–6	Much linear structure (Fig. 3)
(8) Teepol swab, H + C H_2O, ethanol, ethanol vapour	30–40	3–4	4–6	Less structure (Fig. 2)
(9) As (8) + ion bombardment for 1 hour at 200° c	30–40	3–4	4–6	Less structure
(10) As (8) + flamed to red heat	30	3–4	4–6	Less structure
(11) As (8) + flame-melted, held vertically	10	1·5–2·5	3–4	No structure (Fig. 5)
(12) As (8) + flame-melted, horizontal C support	10	—	—	No structure
(13) As (8) + vacuum-melted in Pt tray	10	1·5–2·0	2–3	No structure
(14) Vacuum-melted lead germanate glass	—	1·3–2·0	2–3	No structure (Fig. 6)

Secondly, we observed that a very thin film of zinc evaporated on to the substrate surface was effective in highlighting any inhomogeneities. Zinc was evaporated from a molybdenum boat at a pressure of 10^{-3} mm Hg on to the surface at room temperature. The zinc film was then examined by transmitted and reflected light at various magnifications.

Thirdly, we have used an etching procedure—based on that of Levengood (1959)—to reveal scratch marks and incipient flaws. In this procedure a solution of 6% HF

Further cleaning of the slides by ion bombardment, using a conventional glow discharge in air, with shielded ring electrodes at 2 kv taking about 100 ma, was then tried, the glass substrate being at room temperature or at 200° c. Slides were also baked *in vacuo* to 400° c for several hours or heated in a flame to just below the softening point of the glass.

However, as described below, a marked improvement was observed only when the surface of the glass was melted, as indeed was indicated by Rayleigh (1911, 1912). With the

size of slides used, good results were obtained by holding the slide vertically between two inclined bunsen burners until it became soft, and removing it before excessive distortion occurred. The slide could also be supported horizontally on a heated support, which enabled the upper surface to be flame-melted without distortion, but if heating was too severe some form of thermal etching occurred.

The improvement observed with flame-melted surfaces suggested that complete melting and reforming of the slide was a possibility. To eliminate the evolution of water and gases during vacuum baking, we used glass which had been previously melted in vacuum. Such glass was then melted in a platinum tray, 2 in. × ¼ in. × ⅛ in. deep, heated directly from a low voltage supply. When the surface of the melt was free from bubbles, the current heating the platinum was turned off and the glass solidified. If the amount of glass used was sufficient to completely fill the platinum tray, smooth and flat surfaces were obtained.

Microscope slide glass was vacuum-melted in platinum containers and several substrates were prepared. However, the temperatures involved were inconveniently high and it is not known to what extent constituents of the glass, other than water, were lost during the vacuum melting. Previous work (Fray and Nielsen 1961) had shown that vacuum-melted lead germanate glass (PbO 30, GeO_2 70 at. %) could be prepared more easily and, as this gave similar results, it was used subsequently. (Lead silicate glasses may also be suitable (Adams 1961).)

Results of surface tests

The weight required on the titanium point to give a visible mark on an uncleaned slide was 400 g. Washing the slide with detergent, followed by water or water and ethanol, and allowing it to drain dry, caused a reduction in required weight to 100 g. The weight was further reduced to 30–40 g, if, following the detergent and water wash, the slide was suspended for 10 minutes in ethanol vapour, or was heated in a flame to just below the softening point of tne glass. A further reduction to 10 g was obtained if the surface of the slide was melted in the ways described above. Vacuum-melted germanate glass did not give a scratch mark; cleaved mica, by contrast, required weights of only 1–2 g.

Examination of the zinc film on an uncleaned slide showed an uneven and patchy deposit. In some cases the patches occupied large areas of surface. After the detergent and water wash, the surface was covered much more evenly; certain obvious irregularities could be associated with deposit left by the evaporation of the last traces of solvent used for the final wash. The most even zinc films were produced on surfaces which were blown dry with a jet of clean air (from evaporating liquid air), and on surfaces which had been suspended in ethanol vapour. However, in both these cases irregularities in the film were observed (Fig. 2) and these irregularities were observed also on slides which, following

Fig. 3. Zinc evaporated on slide, finally wiped with Selvyt cloth. (×50)

Fig. 2. Zinc evaporated on vapour-cleaned slides. (×50)

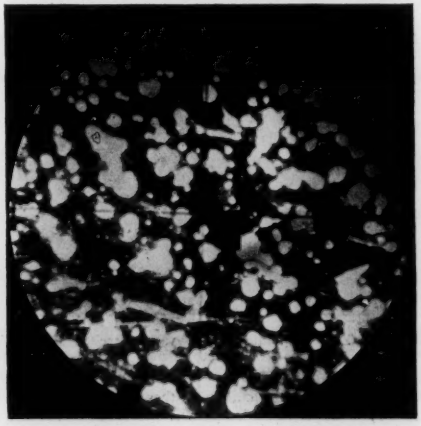


Fig. 4. Zinc evaporated on slide, finally blotted with Selvyt cloth. (×50)

the final ethanol vapour clean, were either ion bombarded for one hour at 200° C or were heated in a flame to temperatures just below the softening point. A most irregular zinc film was obtained if the water or ethanol in the final wash was removed by wiping with a clean Selvyt cloth (Fig. 3) or blotting with a Selvyt cloth (Fig. 4). By contrast, the flame-melted surfaces (Fig. 5) and, especially, the vacuum-melted surfaces (Fig. 6) produced very uniform zinc films.

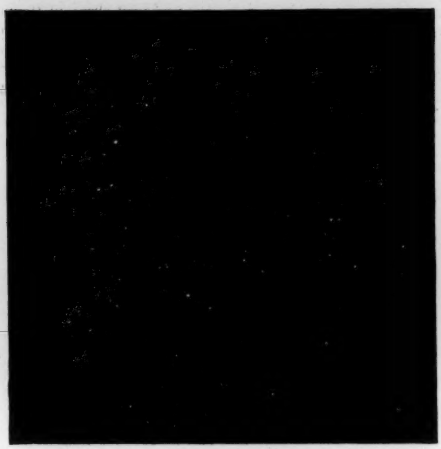


Fig. 5. Zinc evaporated on to flame-melted surface. (×50)

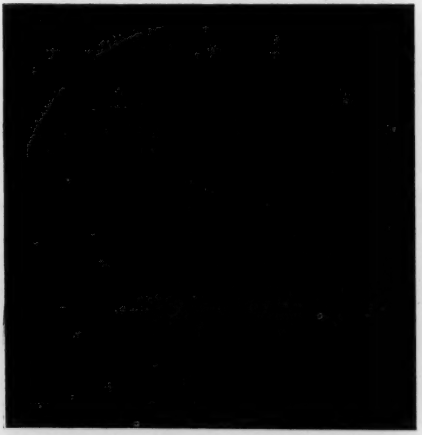


Fig. 6. Zinc evaporated on to vacuum-melted surface. (×50)

It was observed in the course of these experiments that a much higher supersaturation of zinc was required to form a film on glass surfaces which were deliberately contaminated by hydrocarbons, or which had been cleaned by melting.

Typical etch pit patterns are illustrated in Figs 7–10. Fig. 7 shows a scratch on a microscope slide after etching. Figs 7, 9 and 10 show typical features in the middle, at the edge and near the edge of microscope slides. The high etch

Fig. 7. A scratch on a glass slide after etching. (×50)

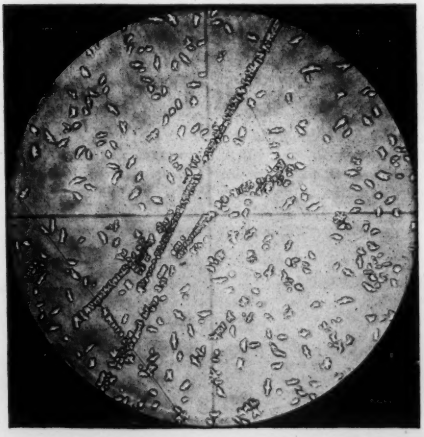


Fig. 8. Typical irregular distribution of etch pits on a slide. (×50)

pit density near the edge of the slide (Fig. 10) was only obtained on some slides. All the slides examined, however, showed an irregular distribution of etch pits. The etch pit patterns from flame-melted slides showed an absence of scratch marks and a rather more regular distribution of etch pits, in general of about the same density as ordinary microscope slides. In those cases where flame melting had distorted the slide, lines of etch emanating from the distorted region (Fig. 11) were often observed. Vacuum-melted micro-

scope slide glass produced fairly uniform patterns with a rather different shape of etch pit (Fig. 12). Vacuum-melted germanate glasses did not appear to etch in the same way with this reagent.

The examination of surfaces using electron microscope replica techniques is at the moment still in a preliminary stage,

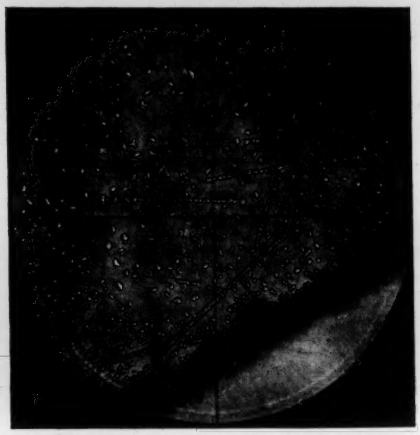


Fig. 9. Typical etch pit pattern near the edge of a slide. (×50)

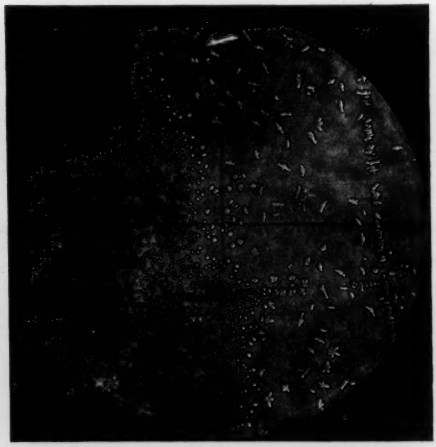


Fig. 10. Etch pit pattern observed near the edge of some slides. (×50)

and has been confined to surfaces produced on the vacuum-melted germanate glasses. Results on the slides examined showed parts apparently free from structure and parts with structure. A rough estimate suggested that about half the surface was free from structure. The main type of structure observed was a 'cobblestone' structure with 'stones' about 70 Å across and about 10 Å high.

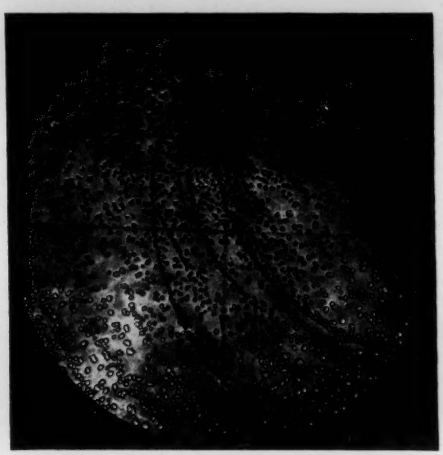


Fig. 11. Etch pit pattern observed near a distorted region of a flame-melted slide. (×50)

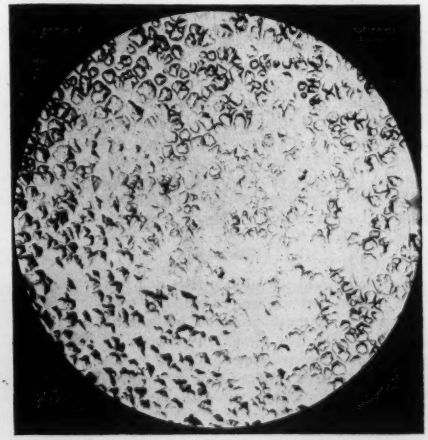


Fig. 12. Typical etch pit pattern on a vacuum-melted slide. (×50)

Permalloy films

Many factors appear to influence the magnetic properties of evaporated Permalloy films. These include composition, rate of deposition, angle of incidence of vapour beam, thickness of film, substrate temperature, impurities, vacuum conditions during evaporation, stress in the film, and the chemical and physical nature of the substrate (cf. Prutton and Bradley 1960 a, b, c). In these experiments we have attempted to

keep all the factors constant except for the physical nature of the substrate. We have altered this by using the cleaning procedures described above. The fact that several slides treated in different ways could be used side by side in the same evaporation helped to eliminate variations between different runs.

The vacuum system was of a conventional type, with a 12 in. glass bell jar pumped by a 2 in. mercury diffusion pump. A 1 g charge of 82% Ni 18% Fe was evaporated from a thick tungsten coil held vertically above the substrate. A copper block substrate heater supported ten variously prepared slides 3 in. × ¼ in. at a temperature of 200° C and an external field of 100 Oe was maintained across the slides. A copper foil shutter between the substrate and coil allowed the coil and charge to be outgassed before the evaporation. After outgassing and gettering with titanium the pressure in the system was approximately 10^{-6} mm Hg and increased to approximately 10^{-5} mm Hg during the evaporation. An evaporation lasting about 30 seconds produced a film of between 1500 and 2000 Å thick.

After cooling to room temperature *in vacuo*, the B/H loops were measured in the 'easy' and 'hard' directions in an apparatus developed by Moore (1960). Square loops were obtained in the 'easy' direction giving a value H_c, representing the domain wall coercivity, and inclined straight lines were obtained in the 'hard' direction giving a value H_k, representing the rotational coercivity. A failure to obtain the square loop and straight line was attributed to defects in the substrate producing anisotropy in directions other than that of the external magnetic field, in the same way as polishing grooves in the substrate create their own anisotropy (Moore and Young 1960). Our results must be regarded as comparative values, as some variation was experienced between runs and as neither the temperature of the substrate nor the composition of the film were determined accurately.

The values of coercivity obtained on variously prepared slides ranged from greater than 8 Oe for the uncleaned slides to less than 2 Oe for the vacuum-melted slides. The Table shows a correlation between magnetic results and the results obtained from the surface tests described previously. In addition the 'squareness' of the loop improved with cleaner surfaces.

Discussion

The results from the surface tests and magnetic measurements suggest that surfaces prepared by melting procedures are superior to those prepared by conventional cleaning techniques. The best surfaces were prepared by quenching molten vacuum-melted glass, and such surfaces—if prepared inside the vacuum system—should be free from dust and absorbed chemicals, free from scratches and polishing marks and should not evolve water or gas during subsequent heat treatment. They may also be expected to be atomically more smooth than the weathered glass surfaces, and the preliminary results from the electron microscopic examination give some support to this.

However, the chemical nature, as well as the cleanliness and smoothness of the surface will be altered by the melting procedure, and the tests used cannot sufficiently differentiate between these.

Acknowledgments

The authors are indebted to Dr. A. C. Moore for loan of equipment to measure B/H loops, to M. C. Hopper and S. Highcock for expert assistance, and to the Controller, H.M. Stationery Office for permission to publish this paper.

References

ADAMS, R. V., 1961, *Phys. Chem. Glass*, **2**, 50.

FRAY, A. F., and NIELSEN, S., 1961, *Infra-Red*, **1**.

GRIFFITHS, A. A., 1920, *Trans Roy. Soc.* A, **221**, 163.

HOLLAND, L., 1958, *Brit. J. Appl. Phys.*, **9**, 410.

HOLLAND, L., and BATEMAN, S. K., 1960, *Brit. J. Appl. Phys.*, **11**, 382.

JOOS, P., 1957, *Z. Angew Phys.*, **9**, 556.

KONIG, H., JÜ LÖFFLER, H., and LAPPE, Fr., 1955, *Glastech. Ber.*, **28**, 131–6.

KORANYI, G., 1959, *J. Soc. Glass. Technol.*, **43**, 438T (1959).

KRUSZEWSKI, S., 1959, *J. Soc. Glass. Technol.*, **43**, 359.

LEVENGOOD, W. C., 1959, *J. Appl. Phys.*, **30**, 378.

MOORE, A. C., 1960, *Inst. Radio Engrs Trans on Component Parts*, CP-7, No. 1, 3.

MOORE, A. C., and YOUNG, A. S., 1960, *J. Appl. Phys.*, Suppl. to Vol. **31**, 1279.

PRUTTON, M., and BRADLEY, E. M., 1960 a, *Proc. Phys. Soc.*, **75**, 557.

—— 1960 b, *J. Appl. Phys.*, Suppl. to Vol. **31** (5), 285 S.

—— 1960 c, *J. I.B.M. Res. and Dev.*, **4** (2).

PUTNER, T., 1959, *Brit. J. Appl. Phys.*, **10**, 332.

RAYLEIGH, LORD, 1911, *Nature, Lond.*, **86**, 416.

—— 1912, *Nature, Lond.*, **90**, 436.

TODD, B. J., 1955, *J. Appl. Phys.*, **26**, 1238.

TSUCHIKASKI, S., SEKIDO, E., and NAKATAM, Y., 1958, *J. Ceram. Ass. Japan*, **66** (754), 233.

YARWOOD, I., 1956, *High Vacuum Technique* (London: Chapman and Hall), p. 152.

High power heating of a cylindrical tube by electron bombardment of the inside surface

by R. E. HAIGH,* M.A., A.M.I.Mech.E., and P. H. DAWSON, B.Sc., Research Department, Associated Electrical Industries (Manchester) Ltd., Trafford Park, Manchester 17

MS. received 5th July 1961

Abstract

When it is necessary for the high heat rating of power reactor fuel elements to be simulated in 'out of pile' tests, the heating of the test element can be a major problem. The method of heating chosen for a particular rig will depend on the exact nature of the tests for which the rig is designed, and the requirements of the tests may make conventional heaters unsuitable. Very high heat fluxes can be obtained by electron bombardment and this method has been used to simulate reactor heat ratings in rig tests on a Calder Hall type element. The test element was essentially a coaxial diode; a vacuum tube representing the fuel rod was the anode, the cathode being a tungsten wire stretched down the centre. Maximum reactor heat ratings were achieved easily and it is considered that if required the power input could have been several times greater. This paper discusses the characteristics of the method and describes the important features of the design.

This paper was presented at the Symposium on the Design of Experimental Equipment for use in Nuclear Energy Research, 11th–14th April, 1961, held at Atomic Energy Research Establishment, Harwell.

1. Introduction

In a Calder Hall type nuclear reactor the fuel element is a uranium rod encased in a thin Magnox can. The heat released by nuclear fission is removed by conduction through the uranium and can to the cooling gas. The radial temperature distribution in the element is indicated in Fig. 1. If the neutron flux is uniform in the region of the element the heat release within the uranium fuel is uniform and this produces a parabolic temperature distribution in the uranium.

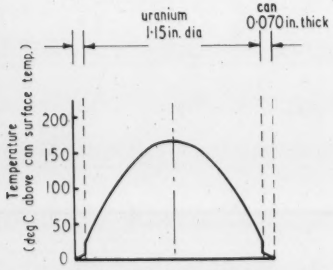


Fig. 1. Temperature distribution in fuel element for a rating of 5 MW/tonne (1·62 kw/in.).

* Now with English Electric Co. Ltd., Atomic Power Division, Whetstone, Leicester.

There is a temperature drop between the uranium and Magnox due to the thermal resistance of the interface and there is also a small drop through the wall of the can.

Many studies of the behaviour of fuel elements in service are more easily carried out by laboratory tests with simulated reactor conditions rather than in a reactor where the conditions cannot be as readily controlled. For tests on the heat transfer performance, the evolution of heat from the fuel by fission must be simulated by heating the element using a non-nuclear method, the choice of a suitable method for a particular test depending on which reactor conditions must be represented exactly. Uniformly distributed heating can be obtained 'out of pile' by passing a heavy current through the uranium. This method, however, can only be used if the required test conditions allow the use of insulation between the uranium and the can. For the actual investigation under consideration—that of the thermal contact resistance between the uranium and can—this was not possible, although the parabolic temperature distribution in the uranium would have been desirable for the correct simulation of the differential thermal expansion of the uranium and can.

An alternative to uniformly distributed heating is to use a uranium tube instead of a rod, and heat the inside of the tube. In this case the radial temperature distribution in the uranium is logarithmic. Fig. 2 shows such a distribution in

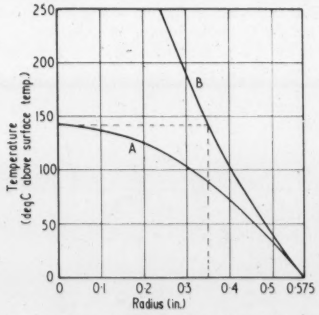


Fig. 2. Temperature distribution in uranium for a rating of 1·62 kw/in. A, parabolic for bar with uniformly distributed heat source; B, logarithmic for tube with separate internal heat source.

a tube 1·15 in. outside diameter, and also the parabolic distribution for a uniformly distributed heat source in a solid rod with the same total heat input. It can be clearly seen that if a hole for a heater in the centre of the uranium were

small, the temperature at the inside surface would have to be high to simulate reactor heat ratings. If the maximum temperature were not to exceed that for uniformly distributed heating, the inside diameter of the tube would have to be not less than 0·7 in.

For the determination of the thermal resistance of the interface, temperature measurements were required using thermocouples within the wall of the uranium tube. Thus a reasonable wall thickness was necessary, and a tube 0·7 in. inside diameter was chosen to represent the fuel rod. A thick-walled tube was also desirable for the radial thermal expansion of tube to approximate to that of a solid rod with uniformly distributed heating. Theoretically, assuming the uranium has constant physical properties and behaves elastically, the overall expansion of a thick tube with a logarithmic radial temperature distribution is not very different from that of a rod with an equivalent parabolic distribution. The expansions are the same if the tube has an inside to outside diameter ratio of 0·53.

For heating the inside of the tube an electron-bombardment technique was chosen as there was considerable doubt whether a conventional resistance type heater, insulated from the uranium, could give sufficiently high heat ratings and a heat input sufficiently uniform axially and circumferentially, to represent reactor conditions. The heater was simply a hot tungsten filament wire stretched along the centre of the tube. With a vacuum inside the tube, a d.c. potential difference between the tube and filament caused the tube to be bombarded with electrons emitted from the filament. The energy to accelerate the electrons was given up as heat as they struck the tube wall. High heat input could be obtained with emission currents of a few amperes by applying about 1000 volts between the tube and filament.

2. Electrical characteristics of the heater

The principles of electron emission are well known (Terman 1943). A hot filament in a vacuum emits electrons which form a cloud round the filament. If there is an electrode called an anode in the region of the filament at a small positive potential with respect to it, the potential field will remove electrons from the cloud to the anode, allowing more electrons to be emitted from the filament. Hence an electric current flows from the anode to the filament. If the potential difference between the anode and filament is increased, the emission current increases until electrons are being removed from the electron cloud as fast as they can be replaced by emission from the filament. At this stage further increase in emission current can only be obtained by raising the temperature of the filament to increase the rate at which electrons are emitted. This condition is known as temperature limited emission. Before the potential field is strong enough to remove electrons as fast as they can be emitted from the filament, the emission current is space charge limited. The power input, which is the product of the applied volts and emission current, is dissipated as heat where the electrons strike the anode.

When using electron bombardment as a method of heating, temperature limited emission is unsatisfactory as it is difficult to keep the emission current constant unless elaborate equipment is used (Allenden 1959). Small changes in filament volts affect the filament temperature resulting in large changes in emission current. Also the emission current depends on the cleanness of the filament and the degree of vacuum. Conversely, with space charge limited emission, the emission current is independent of filament temperature and reasonably independent of the degree of vacuum.

In a coaxial diode where the filament wire is central inside a tubular anode the emission is uniform axially, provided the filament is hot enough along the whole of its length for the emission to be space charge limited, and thus independent of axial variation in filament temperature. It is interesting to note that the emission is also independent of the anode temperature and therefore the heat input to the tube is independent of the tube temperature.

Circumferential uniformity of the heat input depends on the filament being central in the tube. For a small eccentricity of the filament, the local heat input is inversely proportional to the distance between the filament and the tube. Thus for a slightly eccentric filament, the percentage variation in heat input above or below the mean value is equal to the eccentricity expressed as a percentage of the tube radius. For example, a filament eccentricity of 0·01 in. in a tube with a bore 1 in. diameter would produce about $\pm 2\%$ circumferential variation in heat input. Variation in heat input circumferentially tends to be reduced by thermal bowing of the tube. The side with higher heat input heats up more than the other side and bends away from the filament producing a stabilizing effect on the heat input variation.

The space charge limited emission current varies as the applied potential to the power 1·5, and inversely as the tube diameter. Curves of emission current against applied volts are shown in Fig. 3 for various tube diameters with a thin

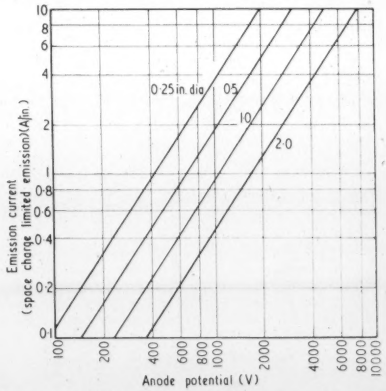


Fig. 3. Relation between emission current and anode potential for various anode diameters.

filament. These curves are only approximate as the emission current varies slightly with the ratio of the filament and tube diameters (Terman 1943), and magnetic effects of the filament heating current also affect the characteristics. Fig. 4 shows the relation between the power input to the tube by electron bombardment and the applied volts for the same range of tube diameters. It can be seen that if a high voltage is used, a power input of several kilowatts per inch can be obtained with an emission current of only a few amperes per inch.

With space charge limited emission there is a limit to the power input which can be reached, and this is governed by the life of the tungsten filament. To achieve enough emission

for high power input, the filament temperature must be in the region 2500–3000° K, and at these temperatures significant evaporation of the tungsten occurs. A higher emission current requires a correspondingly higher filament temperature and this results in an increased evaporation rate. The life of a filament cannot be predicted accurately but a useful indication of the expected life can be inferred from the theoretical time for the loss of a particular proportion of the tungsten by evaporation. Fig. 5 shows how the expected lives of filaments 0·5 mm diameter to 1·25 mm diameter depend on the emission current. These curves are for a 10% reduction in weight based on data given by Jones and Langmuir (1927). Filaments of small diameter have shorter lives than those of larger diameter because they need to run at a higher temperature for the same emission. Filaments with diameters greater than 1·25 mm would have longer lives, but the magnetic fields associated with the much larger filament heating currents required, seriously affect the emission current.

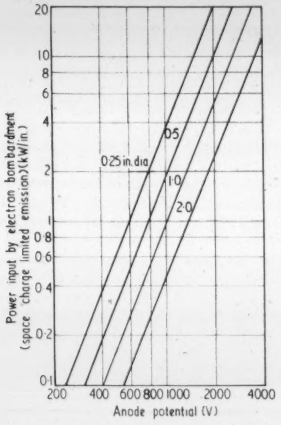


Fig. 4. Relation between power input by electron bombardment and anode potential for various anode diameters.

Figure 5 suggests that a 1·25 mm diameter filament could run for 100 hours with an emission current of 4·5 A/in. With an anode potential of 3000 v in a 1 in. bore tube, the emission would be space charge limited, and it would thus be sufficient for a uniform power rating of 13·5 kw/in. The power rating could be much higher than this for the same emission current and filament life if a higher anode voltage were used. In this case, however, the emission would be temperature limited and the heat input would not necessarily be uniform axially as it would depend considerably on slight axial variations in filament temperature.

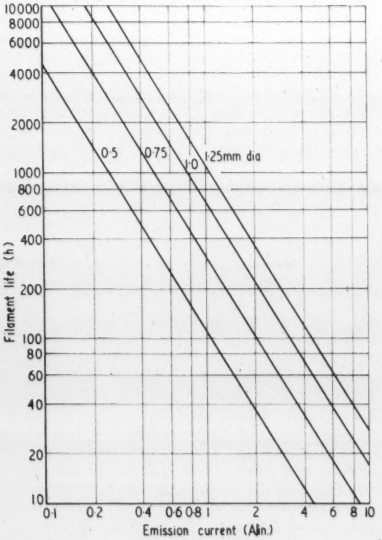


Fig. 5. Relation between the expected life of a tungsten filament and emission current for various filament diameters.

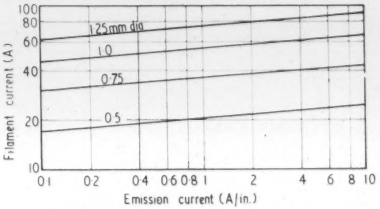


Fig. 6. Currents for heating tungsten filaments of various diameters in relation to the emission current required.

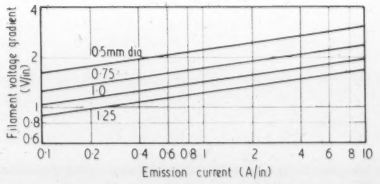


Fig. 7. Voltages for heating tungsten filaments of various diameters in relation to the emission current required.

Alternating currents and voltages required to heat the filament are shown in Figs 6 and 7 plotted against the emission current required. These curves are also based on data given by Jones and Langmuir (1927). Some useful data on the characteristics of tungsten filaments are also given by Kohl (1951). The energy to heat the filament is dissipated by radiation and this results in extra heat input to the tube wall. The total power input is therefore the sum of the power supplied by electron bombardment and about 0·1 kw/in. supplied by filament heating.

The emission current also heats the filament as this current has to flow along the filament to the terminal. It is however a direct current superimposed on the alternating filament heating current and the resulting total r.m.s. current is very little higher than the r.m.s. value of the alternating filament current. In fact the heating effects of the two currents can be considered separately and if the square of the total

emission current is small compared with the square of the filament current, then the emission current filament heating effect is negligible.

3. Design and operation of the test rig

The general arrangement of the test section for interface resistance measurements using a specimen heated by electron bombardment is shown in Fig. 8. The fuel element was

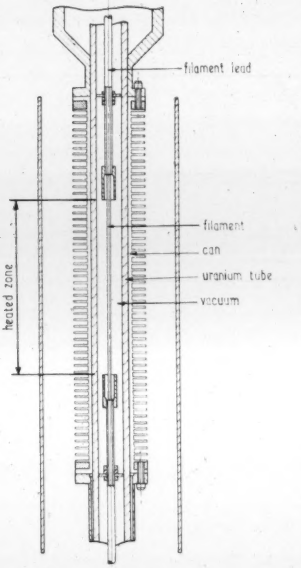


Fig. 8. Arrangement of test section interface rig.

located in a test chamber forming part of a pressurized CO_2 circuit with the specimen vertical so that it was easier to arrange for the filament to be central inside the tube. Although most tests were carried out with a steel tube, a uranium tube was used for some tests and the heater operated satisfactorily in both cases. The specimen, consisting of the tube with a Magnox can pressurized on to it, was assembled in the rig by connecting extensions to the ends of the tube and can, and the whole test assembly was supported from the top flange of the test chamber so that it could be removed from the rest of the rig as a unit.

As the heater was essentially a high power co-axial diode, many of the design details were based on experience gained in the design of high power valves. During operation the inside of the specimen was continuously pumped with an oil diffusion pump. So that a good vacuum could be maintained all the materials of the vacuum circuit had a low vapour pressure at operating temperature and where necessary components were out-gassed and checked for porosity during manufacture. Constrictions impeding evacuation were avoided as much as possible. Several types of vacuum seals were used. At operating temperatures up to 150° C, silver-plated steel 'O'-rings, and copper 'O'-rings smeared with silicone varnish, were found to be satisfactory; other types such as rubber 'O'-rings were used where the temperature was low or where water-cooling could easily be arranged. Although they have not yet been used in the rig, it is believed that silicone rubber 'O'-rings will operate at 200° C for a few hundred hours.

The specimen tube was bored straight from stress-relieved material so that when heated relief of residual stress would not cause bending which would result in non-uniform heating. The ends of the 1·25 mm diameter tungsten filament were atomic hydrogen welded to $\frac{1}{8}$ in. diameter molybdenum rods leading to terminals at the ends of the test section. The joints were prepared for welding by inserting the tungsten wire into short holes drilled in the ends of the rods. The terminals were insulated from the rest of the test section by glass seals fused to 'Nilo-K' rings having a similar thermal expansion coefficient, the 'Nilo-K' rings being joined to the steel components of the rig by copper brazing. Electrical connection to the lower terminal was made with a copper rope covered with an insulating glass sleeve; this passed through a tube leading out through the top flange of the test chamber.

The tensioning device to keep the filament straight originally consisted of a sliding terminal with a bellows seal. This allowed check measurements of the extension of the filament to be taken during tests so that the tension on the filament could be adjusted as required. This system needed very careful operation and plastic extension of the filament sometimes occurred because it was not possible to correct the filament tension quickly enough with changes in filament temperature. A later modification was made so that a fixed tension of about $\frac{1}{2}$ lb was applied to the filament by a light spring inside the vacuum chamber. Although the extension of the filament could not be checked and the tension could not be adjusted during testing, the modified system worked perfectly, keeping the filament straight without straining it plastically.

Insulators fixed near the ends of the filament leads held the filament central in the tube. These insulators were triangles cut from alumina disks. There was a central hole in the insulators so that they could be clamped with molybdenum locknuts on the threaded ends of the filament leads. The corners of the insulators were ground concentric with the hole to make them a sliding fit in the specimen tube. The triangular shape allowed sufficient gap past the insulators for a good vacuum to be maintained in the heater.

The ends of the filament, where the temperature was likely to be too low for space charge limited emission, were masked for a length of $\frac{3}{4}$ in. with thin molybdenum tubes $\frac{5}{16}$ in. diameter fixed to the ends of the filmanent leads. This ensured that nearly all the emission was from the part of the filament between the masks where it was hot enough for space charge limited emission. Thus the total power input by emission could be taken as uniformly distributed along the effective length of the filament.

A simplified diagram of the electric circuit is shown in Fig. 9. The actual circuit was more complicated as the heater circuit was only part of the control circuit for the whole rig including the blower, gas heaters, vacuum pumps, and gauges and many safety devices were incorporated so that the supply to the heater was automatically shut off if a fault occurred in other parts of the rig.

The filament heater supply was 80 A at 25 V. The voltage

required was greater than that indicated in Fig. 7 as there was a significant voltage drop in the filament leads and in the connections between the transformer and the rig. Fig. 9

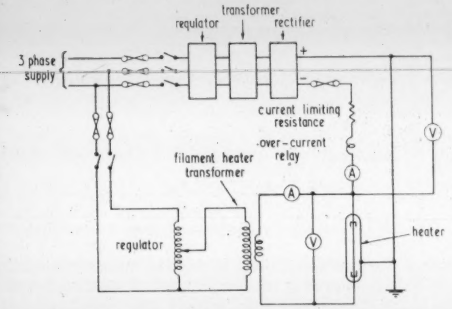


Fig. 9. Electron bombardment heater control circuit.

shows that the whole of the filament circuit operated at a high d.c. potential relative to earth, and the insulation, including the insulation of the secondary winding of the filament heater transformer, had to be sufficient to withstand this potential.

The total d.c. supply available was 15 A at 1500 V, although all of this power could not be used to heat the specimen as a certain amount of power was dissipated in the current limiting resistance, reducing the available voltage at the heater to about 1000 V. The function of this resistance was to limit the discharge current in the heater until the over-current relay had time to shut off the supply if the pressure in the heater rose allowing a discharge to occur. This happened quite often in preliminary tests until the vacuum circuit had been properly outgassed as raising the temperature caused gassing and loss of vacuum. During normal operation the pressure was usually between 10^{-6} and 10^{-5} mm Hg measured in the vacuum manifold.

When switching on the filament heater the voltage supply was increased slowly from zero using the regulator, because the cold resistance of tungsten is of the order of one tenth that at operating temperature, and the current surge from switching a cold filament on to full voltage would damage the filament and probably the transformer. The usual procedure when starting up the emission heating was to set the d.c. voltage between the filament and earth to a suitable value for the power rating required and then increase the filament current until no further increase in emission current could be effected by an increase in filament current. This ensured that conditions of space charge limited emission had been reached.

The maximum uniform power input obtained was a little over 1·6 kw/in. If more power supply were available much higher power ratings could easily be achieved, assuming that sufficient cooling could be arranged to prevent overheating of the specimen.

4. Conclusion

The electron-bombardment method of heating has been used successfully to simulate reactor power ratings in heat transfer tests on a dummy fuel element. Where the method can be used, it has the advantage over more conventional methods of heating in that the power ratings that can be achieved are very much higher. The techniques required are not new as they have already been established in the manufacture of high power valves and high vacuum systems.

Acknowledgments

The authors wish to thank their colleague, Mr. J. H. Ludlow, for much valuable advice. Acknowledgments are also due to The Nuclear Power Group who sponsored the work, and to Sir Willis Jackson, Director of Research and Education, Associated Electrical Industries (Manchester) Ltd., for permission to publish this paper.

References

ALLENDEN, D., 1959, *J. Sci. Instrum.*, **36**, 66.

JONES, H. A., and LANGMUIR, I., 1927, *Gen. Elect. Rev.*, **30**, 310, 354.

KOHL, W. H., 1951, *Materials Technology for Electron Tubes*, (New York: Reinhold).

TERMAN, F. E., 1943, *Radio Engineers' Handbook* (New York and London: McGraw-Hill).

Recording micro-densitometer employing beta-radiation from a carbon-14 radioisotope source

by R. A. LLOYD M.Sc., A.Inst.P., and E. H. ADAMS, The National Chemical Laboratory Isotope Applications Unit, Wantage Research Laboratory, Wantage, Berks.

MS. received 28th July 1961

Abstract

The variation of wood density through the thickness of the individual growth rings of trees is of interest to those concerned with research on the growing of trees and the subsequent utilization of the timber. An experimental instrument has been developed in which a finely collimated beam of carbon-14 beta-particles is used to measure the density variations along thin longitudinal sections cut with a microtome from radial core borings obtained from the trunks of standing trees. The response of the instrument is linear for densities up to at least 1·0 g/cm³ with sections 100 microns thick, and density discontinuities 0·05 mm apart can be resolved. Density changes of 3·5% can be detected, at a speed of traverse of 0·5 mm/min. The instrument may well be applied to other materials.

1. Introduction

ALTHOUGH the instrument is suitable for measuring, with good resolution, the variation of mass per unit area of any thin material, it was designed with a specific application in view. In this application it is used to measure the density distribution along the length of sections of wood 8 in. or more × ½ in. × 100 μ thick, as distinct from previous measurements on a semi-microscale using thick specimens (Cameron *et al.* 1959). These sections were cut with the aid of a microtome from the cores of radial borings which traverse the growth rings of standing trees. The density distribution is therefore obtained without destroying the tree, and the information can be used as a guide to sylvicultural treatment, for assessing the value of trees, and for selecting those required for propagation. Details of the use of the instrument in this context will be the subject of a separate communication to be published by the Forest Products Research Laboratory, Princes Risborough, on behalf of whom the instrument was designed.

2. Factors affecting the design

Two criteria which had to be satisfied were an adequate geometrical resolution and an adequate sensitivity. The geometrical resolution is defined as the least distance separating two thin wires, representing density discontinuities, at which they can be distinguished. It will be seen that this corresponds to the separation of two over-lapping response peaks, which is a familiar problem. The second criterion, the sensitivity, is defined as the smallest change in density which can be detected. This is the same as the smallest change in count-rate which is significant of a change in density, and therefore will obviously depend on such factors as the time of observation, the background and the count-rate.

The system used was a radiation source, a slit assembly and a detector. The factors determining the choice of the radiation source were (i) that the radiation should be emitted by an isotope of long half-life, (ii) that it should have an energy which gave the desired beam attenuation over the range of densities to be encountered (the optimum condition is that 87% of the radiation is absorbed by the specimen), and (iii) that the source of radiation should be available at a very high specific activity, so that collimation can be used to obtain good resolution with the minimum loss of beam intensity. The source chosen was carbon-14 in crude lipid material from the alga *chlorella*-C14, which provides one of the highest specific activities available for this isotope.

The detector was required to have a high efficiency for the detection of low energy beta-particles, to be stable in operation, and to be convenient to use. It would also be required to operate behind one of the slits forming the collimation system for the radiation beam. A scintillation counter was chosen because there is some flexibility in the choice of scintillator and its positioning with respect to the photomultiplier and slit. A thin disk of anthracene cemented to a light guide was finally used.

3. Description of the instrument

The details of the apparatus are shown schematically in Fig. 1. The specimen to be examined is kept rigid and flat

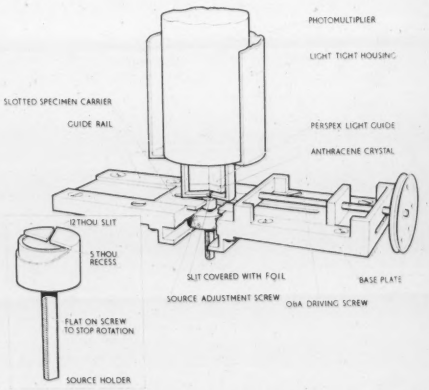


Fig. 1. Details of the apparatus showing the specimen carrier in position and the source raised into the specimen carrier slot.

by sticking it to the slotted specimen holder. This slides between the guide rails on the base plate. A hole in the centre of the base plate accommodates the source, which can be raised to enter the slot in the specimen holder. This is necessary because the distance between the source and the detector must be kept as small as possible so as to achieve the maximum counting rate. A 0 B.A. driving screw, 1 mm travel per turn, is the traversing mechanism by which the specimen is moved between the source and the detector. The screw is driven at constant speed by a synchronous electric motor, but the rate of movement of the specimen may be altered by changing the pulley wheel ratio.

The scintillation detector consists of a light-tight housing containing the photomultiplier and Perspex light guide, to which is cemented a 1 mm thick anthracene disk. A hole machined in the end of the housing was covered on one side with a selected piece of aluminium leaf 30×10^{-6} in. thick, and on the other side with two razor edges which formed one slit and which could be adjusted for varying the slit width.

A very high intensity beam of carbon-14 beta-particles was required from the source, and the design shown in Fig. 1 was eventually adopted. The brass cap $\frac{1}{16}$ in. thick had a slit 0·3 mm wide cut across a diameter of its end face. The brass post, which was a sliding fit in the cap, accommodated the source in a recess. The source was deposited in this recess, covered with aluminium leaf, and then protected by cementing the brass cap in position.

A threaded shaft attached to the source allowed it to be raised and lowered, but a flat was cut on one side of the shaft to prevent it from rotating and throwing the slits out of alignment. The alignment of the source and detector slits was critical, and necessitated a mounting which was adjustable but which could be rigidly locked in position.

The associated electronic apparatus for detecting, amplifying and recording the pulses was conventional except that provision was made for feeding a 2-decade logarithmic ratemeter from the Dekatrons of a scaler. This gave a good scale length on the strip chart recorder for ease in reading, as well as accurate range changes by factors of 10, by switching the ratemeter to the output of the appropriate Dekatron. As the input to the logarithmic ratemeter depends exponentially upon the wood density, the overall response of the instrument within its working range is sensibly linear.

4. Performance

4.1. *The detector*

The optimum working conditions for the scintillation detector were found by plotting the e.h.t. characteristics. An anthracene disk cut from a single crystal was compared with a plastic scintillator disk of similar size, and curves for source and background were plotted. It was clear that anthracene provided a better discrimination between source and background than did the plastic.

4.2. *Alignment of the slits*

The count-rate is a maximum when the source and detector slits are accurately aligned. Fig. 2 shows the variation in count-rate as one slit is moved across the other. A displacement of as little as 0·025 mm could cause a 10% change in the counting rate. An indication of correct alignment can also be obtained by observing the counting rate as the two slits approach each other, as shown in Fig. 3. The curve should not have a maximum point which would indicate that one slit is offset from the other.

The correct angular alignment of the slits can be made in a similar way, as shown in Fig. 4. The central position was in this case found optically, but normally this would not be done because it involves dismantling the detector. The displacement of the peak from the central position was due to a small offset of the slits during the measurement. This was corrected before further measurements were made.

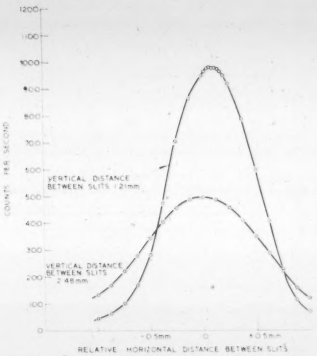

VERTICAL DISTANCE BETWEEN SLITS 1·21mm
VERTICAL DISTANCE BETWEEN SLITS 2·46mm
RELATIVE HORIZONTAL DISTANCE BETWEEN SLITS

Fig. 2. The variation in counting rate as one slit is moved across the other, but with a constant vertical distance between them.

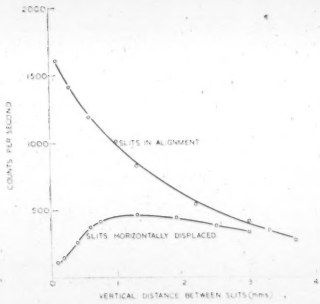


Fig. 3. The variation in counting rate as the vertical distance between the slits is varied.

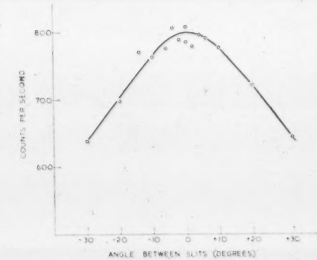


Fig. 4. The variation in counting rate as the angle between the slits is varied.

4.3. *The density calibration*

The calibration for density was made with a step-wedge composed of layers of nylon film 0·01 mm thick. The strip of film was selected for uniformity by measuring the variation in mass per unit area of many different pieces. A uniform piece was then calibrated by measuring its thickness and by weighing a known area. Fig. 5 shows the calibration curve,

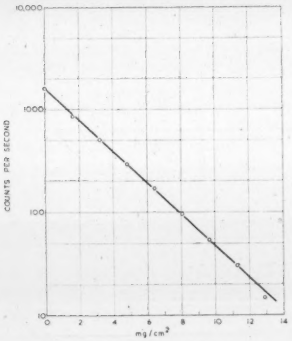


Fig. 5. The density calibration curve.

which is linear up to about 12 mg/cm^2, which is almost the limit of the useful range of the instrument because of the approach to background at this end of the scale. The working range is dependent upon the width of the slit, which in turn governs the resolution. The slit width was later reduced to improve the resolution to 0·05 mm (50 μ) and this reduced the upper limit of the working range to 8 mg/cm^2. The different absorption coefficients of wood and nylon arising from differences in hydrogen content do not significantly affect the shape of the curve.

4.4. *The response time*

If a straight edge is moved between the two slits, the graph of counting rate plotted against distance gives a response curve for the instrument as shown in Fig. 6. This curve

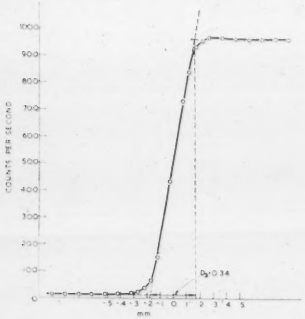


Fig. 6. The change in response as a straight edge passes between the slits.

should be symmetrical in order to avoid distortion of the density distribution as plotted by the recorder. The shape of the curve depends on the time constant of the recording system and on the speed at which the edge moves between the slits. For a time constant of 15 seconds, the rate of travel could not be faster than 0·3 mm/min to achieve a symmetrical response.

4.5. *The geometrical resolution and response curves*

Although the resolution is governed by the slit width, it is not necessarily equal to it because of factors such as beta-particle scattering, source inhomogeneity and slit alignment. It is therefore necessary that it should be checked experimentally. The geometrical resolution has already been defined as the least distance which must separate two thin wires in order for them to be distinguished. Fig. 7 shows

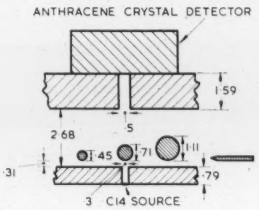


Fig. 7. The arrangement for measuring the response curves of wires. (Dimensions in mm.)

three such wires of different diameters and their positioning relative to the two slits. The response curves are shown in Fig. 8 and it can be seen that each wire gives a response 'peak'.

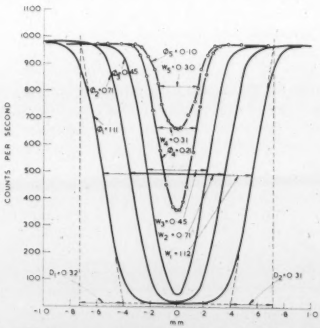


Fig. 8. The response curves for wires of various diameters as they pass between the slits.

Therefore in order for the wires to be recorded separately, their peaks must be recorded separately, and in accordance with usual practice the simplest assumption is that two equal peaks will be resolved provided they do not intersect at a point higher than half their peak height. Since the peaks are identical in shape, one can then measure the resolution by measuring the width of the peak at half the peak height.

If Fig. 8 then is examined in the light of this re-phrased definition of resolution, it can be seen that the half-peak width W depends on the wire diameter ϕ. For the larger

wires, W is equal to ϕ as would be expected from simple geometrical considerations. For the thinner wires, W is independent of ϕ and is the true resolution. In these measurements it was 0·31 mm because the controlling slit was the one over the source.

The significance of the term 'thin wires' in the definition of resolution can now be appreciated; unless the wire is narrower than the slit, the half-peak width will be the wire diameter and not the slit width, and therefore not the resolution. Rigorously, the resolution R is equal to W only when ϕ is less than R.

A further point to note in this connection is that the curve for $\phi = 0{\cdot}45$ mm has a half-peak width equal to this value of ϕ, and yet the curve never reaches background level. The half-peak width confirms the fact that the wire is wider than the slit, and yet the fact that it never reaches background level means that it never completely obscures the radiation beam from the slit. This discrepancy is probably due to beta-particles being scattered into the detector slit from the air surrounding the wire. It also follows that the correct procedure for measuring the resolution is to reduce ϕ until W remains constant. The alternative procedure, to reduce ϕ until the curve fails to reach background level, may give an erroneous result because of the effect of beta-particle scattering.

Lastly, in the curve for $\phi = 1{\cdot}11$ mm the radiation beam has been totally obscured, and therefore each half of the curve is a straight edge response. The width D of the straight edge curve has been indicated on the diagram by a self-explanatory geometrical construction. It can be shown that D should be equal to the width of the slit. In fact $D_1 = 0{\cdot}31$ mm and $D_2 = 0{\cdot}32$ mm, which agrees reasonably well with the width of the source slit (0·30 mm). The width D_3 (Fig. 6) is not in such good agreement (0·34 mm) and this again is probably due to the difference in beta-particle scattering between the sharp edge of the razor blade and the cylindrical edge of the wire. There is some stray radiation escaping laterally from the source slit which would not normally enter the detector slit. The approaching edge can scatter this stray radiation into the detector slit, but the amount of scatter will depend on the shape of the edge.

The best resolution obtainable was 0·05 mm (50 μ) and this was achieved by adjusting the detector slit to this value. A reduction in slit width moves the density calibration curve parallel to itself towards the origin. The upper limit of the density range is therefore reduced to 8 mg/cm^2 under these conditions.

5. Sensitivity

The recorder trace deviates from a steady reading, even when the specimen is stationary, because of the statistical fluctuation in the counting rate. The sensitivity, defined as the smallest change in density which can be detected, would therefore depend on these fluctuations.

If the mean counting rate is N, the output of the ratemeter will have a relative standard deviation equal to $(2NT)^{-1/2}$ where T is the ratemeter time constant. Now it has already been shown from the density calibration curve that the absorption of beta-particles is exponentially dependent upon the density of the specimen, i.e. $N = N_0 \exp(-mpb)$ where m is the mass absorption coefficient, ρ is the density and b, the thickness of the specimen, is constant. N_0 is a constant determined by the strength of the source, the efficiency of the detector and the area of the slit, and is equal to N, the counting rate, when $b = 0$.

Differentiating,

$$\frac{dN}{N} = -mpb\frac{d\rho}{\rho}.$$

This means that for a change in density from ρ to $\rho + d\rho$ the counting rate changes a proportional amount. Let this change in counting rate be equal to three times the standard deviation of the ratemeter output. Such a statistical fluctuation of the counting rate rarely occurs in practice (probability 3×10^{-3}) and therefore such a deviation can be regarded as a significant change of density.

Hence

$$-\frac{d\rho}{\rho} = 3(mpb)^{-1}(2NT)^{-1/2}$$

and replacing mpb,

$$-\frac{d\rho}{\rho} = 3\left(\ln\frac{N_0}{N}\right)^{-1}(2NT)^{-1/2}.$$

This expression for the sensitivity shows that it is not constant as ρ (and hence N) varies. There is a minimum at $N/N_0 = 0{\cdot}135$ (corresponding to $mpb = 2$) but the curve of sensitivity against ρ is relatively flat over a substantial portion of the range. For instance, for a ratemeter time constant of 10 seconds, the best sensitivity occurs at 6 mg/cm^2, where the sensitivity is $\pm 2{\cdot}3\%$, and over the range 2–12 mg/cm^2 the sensitivity is not worse than $\pm 3{\cdot}3\%$. It can therefore be concluded that a deviation of $3\frac{1}{2}\%$ over the range 2–12 mg/cm^2 constitutes a significant change in density.

Acknowledgments

The authors gratefully acknowledge the help of their colleagues, particularly Dr. E. W. J. Phillips of the Forest Products Research Laboratory, who initiated previous related work (Cameron *et al.* 1959), Mr. R. K. Barnes of the Department of Scientific and Industrial Research Isotope Applications Unit who suggested the use of radiation from carbon-14, and Mr. R. F. S. Hearmon of the Forest Products Research Laboratory for valuable discussions.

This paper is published by permission of the Department of Scientific and Industrial Research.

Reference

CAMERON, J. F., BERRY, P. F., and PHILLIPS, E. W. J., 1959, *Holzforschung*, **13**, 3.

Light transmission technique for following continuous changes in optical retardation in polymer films

by J. H. MAGILL, Ph.D., D.I.C., A.R.I.C., Research Department, British Nylon Spinners Ltd., Pontypool, Monmouthshire

MS. received 18*th May* 1961

Abstract

The light intensities (crossed and parallel polaroids) transmitted by a spherulitic polymer film of uniform thickness have been measured using a photomultiplier-recorder system. These intensity measurements have been correlated with the optical retardations (Berek compensator) measured over the same temperature interval. The linear relationship obtained establishes the validity of the technique for following changes in optical retardation. Typical results are given for polyhexamethylene adipamide (66 nylon) and isotactic polypropylene polymer films.

1. Introduction

Several techniques (Hartshorne and Stuart 1960) are available for measuring optical retardation in anisotropic systems but these methods are not particularly suitable for following rapid and continuous changes. During earlier observations made while studying crystallization rates in polymers (Magill 1960) it was found that the change in intensity of light transmitted by a birefringent spherulitic film between crossed polaroids followed a trend similar to the birefringence dependence over the same temperature range (Cannon and Chappel, to be published*). A more detailed study revealed that the technique described for rate measurements (Magill 1961) could be adopted for following continuous changes in overall retardation or birefringence with temperature.

2. Theory

Assuming that absorption, reflection and scattering are negligible, the intensity of light transmitted by a uniformly thick optically anisotropic crystal plate or film (Partington 1953) with crossed polaroids, is given by

$$I_c = a^2 \sin^2 2\alpha \sin^2 (\pi R/\lambda) \tag{1}$$

where a is the amplitude, α is the angle between polarizer and the axis of the crystal, R the optical retardation and λ the wavelength of the monochromatic light.

With parallel polaroids, the intensity is

$$I_p = a^2\{1 - \sin^2 2\alpha \sin^2 (\pi R/\lambda)\} \tag{2}$$

whence it follows that

$$\frac{I_c}{I_0} = \sin^2 2\alpha \sin^2 \left(\frac{\pi R}{\lambda}\right) \tag{3}$$

where $I_0 = I_c + I_p$. This expression eliminates effects of any variation in source intensity, photomultiplier gain or sensitivity and minimizes stray radiation signals.

Spherulites may be considered to have a radially symmetrical distribution of optically anisotropic crystallites. For a simple two dimensional spherulite, the contribution of all orientations to the transmitted light intensity is given by

$$8\int_0^{\frac{1}{4}\pi} \sin^2 2\alpha d\alpha \tag{4}$$

where $0 < \alpha < \frac{1}{4}\pi$. For such a spherulite

$$\frac{I_c}{I_0} = \sin^2 \left(\frac{\pi R}{\lambda}\right) . 8\int_0^{\frac{1}{4}\pi} \sin^2 2\alpha d\alpha; \tag{5}$$

hence

$$\frac{I_c}{I_0} = \pi \sin^2 \left(\frac{\pi R}{\lambda}\right). \tag{6}$$

In general,

$$\frac{I_c}{I_0} = k \sin^2 \left(\frac{\pi R}{\lambda}\right) \tag{7}$$

or

$$\sin^{-1} \left(\frac{I_c}{I_0}\right)^{1/2} \propto R \tag{8}$$

for a uniform spherulitic film for the three dimensional model.

3. Experimental method

An annealed birefringent polymer film was mounted between No. 1 glass cover slips using MS 550 silicone oil as an immersion medium which gives protection against oxidation. The sample was placed on the Köfler hot stage mounted on the polarizing microscope and focused visually. The entire field was filled with spherulites. The microscope substage was removed to reduce light scatter and to facilitate the illumination of the sample with parallel light from a stabilized tungsten source. Light of wavelength 430–490 mμ, selected using an Ilford spectrum blue filter, was used to obtain maximum sensitivity with a Mazda 27 M1 photomultiplier. The photomultiplier was suitably placed over the microscope eyepiece to receive the transmitted light on the photocathode. Changes in the photomultiplier current corresponding to the transmitted light intensity variations were followed using a high-speed Honeywell–Brown millivolt recorder. The recorder response time was 0·25 seconds and full scale deflection, using a suitable load resistance, corresponded to a photocurrent of 10 μA.

Recordings of the transmitted light intensity with crossed polaroids (I_c) together with periodic recordings of intensity with parallel polaroids (I_p), were made by moving the microscope analyser quickly from the crossed to the parallel position and back again during the heating or cooling cycles. A small variation in photomultiplier sensitivity with the plane of polarization was observed, but it was found that the photocathode sensitivity was the same for the two planes of polarization corresponding to I_c and I_p. The zero position was checked periodically during measurements by momentarily shutting off the incident light. The temperature of the sample was continuously varied by manual adjustment

* A summary of this work is published in the report of the Bristol Conference on Physics of Polymers, January 1961 (*Brit. J. Appl. Phys.*, **12**, 261).

of the hot stage rheostat controller. The hot stage mercury-in-glass thermometer temperatures during a run were periodically marked on the recording chart. A record was thus made of changes in the transmitted light intensity during heating and cooling conditions. Good reproducibility was obtained for individual runs on a given sample for both rising and falling temperatures.

Optical retardation measurements were also made, using a Berek compensator, over the same temperature range as the light transmission readings. The mean of several spherulites in the field of view was found at every temperature for each film. These average retardation values were used subsequently throughout this work.

4. Results and discussion

Equations (7) and (8) were verified experimentally using point determinations of I_c and I_p interpolated from the recorder traces. Typical results are illustrated in Figs 1, 2

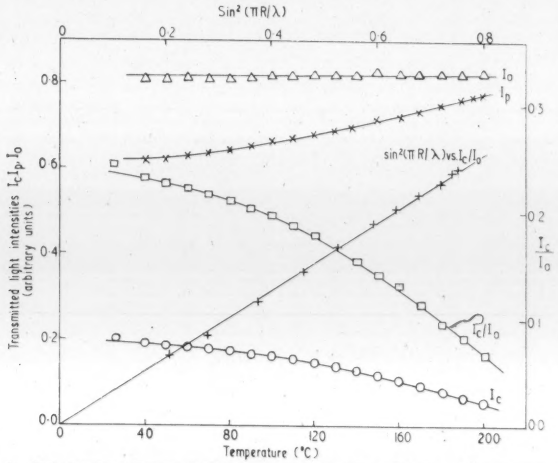


Fig. 1. Relationship between light transmission, optical retardation and temperature for positively birefringent spherulites in annealed 66 nylon film.

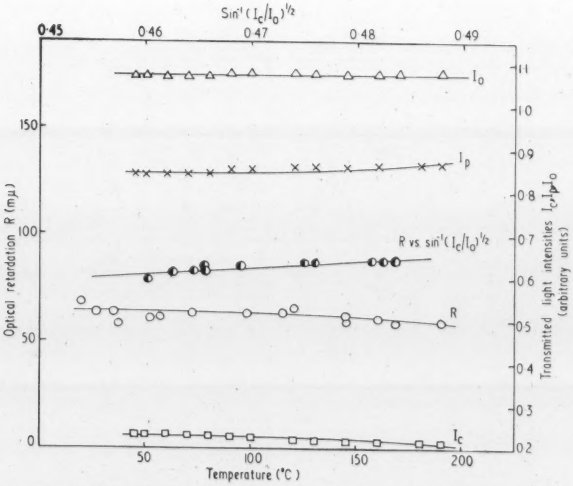


Fig. 2. Relationship between light transmission, optical retardation and temperature for negatively birefringent spherulites in annealed 66 nylon film.

and 3. The graphs of I_c plotted against temperature decrease with increase in temperature in contrast to those of I_p over the same range. These curves are mirror images of each other. For each polymer film examined $I_p + I_c$ was constant ($=I_0$). I_0 was therefore used for normalization purposes.

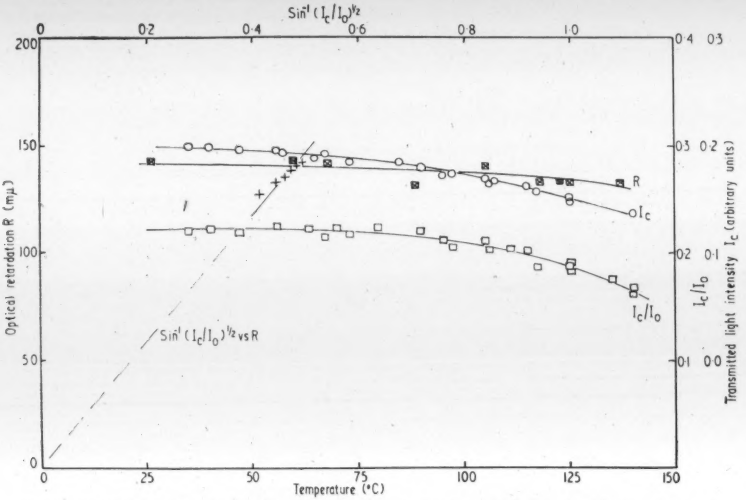


Fig. 3. Relationship between light transmission, optical retardation and temperature for positively birefringent spherulites in annealed isotactic polypropylene film.

Assuming a mean λ of 460 mμ, the graph of I_c/I_0 against $\sin^2(\pi R/\lambda)$ for a positively birefringent spherulitic 66 nylon film is linear as predicted by Eqn (7). Straight line relationships were also found for graphs of $\sin^{-1}(I_c/I_0)^{1/2}$ plotted against R for a negatively birefringent spherulitic 66 nylon film (Fig. 2), and a positively birefringent isotactic polypropylene specimen (Fig. 3). The largest changes in light intensity with temperature were observed for positively birefringent spherulites of 66 nylon. Comparatively small changes were obtained for the other two spherulitic films.

The optical retardation (Berek compensator) measurements show similar trends over corresponding temperature ranges. These results are in agreement with measurements of Cannon and Chappel (to be published) who have made a detailed study of the birefringence of both positively and negatively birefringent 66 nylon spherulites at temperatures up to the melting point of the polymer. The trend observed in isotactic polypropylene is similar to that mentioned by Keith and Padden (1959).

Annealed polymer films were used throughout these experiments because secondary ordering processes especially in isotactic polypropylene produce marked reductions in optical retardation (Keith and Padden 1959) and transmitted light intensity during thermal relaxation cycling prior to the attainment of an equilibrium state. The highest point in the temperature cycle should always be well below the polymer thermodynamic melting temperature T_m to avoid hysteresis effects arising because of appreciable mobility in the solid below T_m.

During the experimental procedure in this paper, care was taken not to exceed a retardation of 230 mμ (λ in these experiments taken as 460 mμ). The transmitted light intensity measurements which were always made on fields of view completely filled by spherulitic regions were unaffected by the shape of the spherulite boundaries. Expected differences in light intensity transmitted by irregularly shaped or unsymmetrically disposed spherulites are presumably averaged when a large number of spherulites are observed.

5. Conclusions

The transmitted light intensity, with crossed and parallel polaroids, can be formulated to give a measure of the optical retardation and hence birefringence of thin spherulitic polymer films of uniform thickness. The linearity of the optical retardations (measured by a Berek compensator) plotted against the normalized transmitted light intensities expressed by theory (Eqns (7) and (8)), establishes the validity of the technique, which can, therefore, be used to follow continuous and rapid changes in optical retardation or birefringence.

6. Acknowledgment

The author is indebted to colleagues at British Nylon Spinners Ltd., for helpful discussion.

References

HARTSHORNE, N. H., and STUART, A., 1960, *Crystals and the Polarizing Microscope* 3rd edn (London: Edward Arnold), Chapter 7.

KEITH, H. D., and PADDEN, F. J., 1959, *J. Appl. Phys.*, **30**, 1479.

MAGILL, J. H., 1960, *Nature, Lond.*, **187**, 770.

—— 1961, *Polymer*, **2**, 221.

PARTINGTON, J. P., 1953, *An Advanced Treatise on Physical Chemistry* (London: Longmans, Green), p. 183.

Dynamic compression measurements of cushioning materials utilizing an analogue computer

by B. H. VENNING, B.Sc.(Eng.), A.C.G.I., A.M.I.E.E., and J. GUTTRIDGE, B.Sc.(Eng.)*, Department of Electronics, University of Southampton

MS. received 17*th May* 1961

Abstract

Theoretical analysis of the design of packaged systems is readily carried out on an electronic analogue computer, but to introduce a realistic stiffness term it is necessary to know the dynamic performance under conditions similar to that of an actual impact. A system has been devised using an electro-dynamic force generator with force and displacement transducers to incorporate a sample of cushioning material into the analogue system, thus solving the cushioning problem under dynamic conditions. At the same time, waveforms of force–time and force–displacement are produced, showing the nature of the non-linear stiffness and the extent of energy dissipation.

1. Compression characteristics of bulk materials

THE transport of fragile equipment usually necessitates careful packaging to avoid damage being caused by the various hazards encountered during movement. The severity of mechanical shock is reduced by incorporating some cushioning material between the packaged article and the outer container in which it travels. Materials such as wood or paper shavings or corrugated cardboard can be used, with the quantity and method used based entirely on experience, with little attempt at obtaining specific or repeatable results. Many applications, however, require packs to be designed to meet firm specifications of performance figures, and most of these employ cushions of some type between the inner and outer containers. Typical bulk cushioning materials at present in use include (Schuler 1958): (*a*) rubberized hair, latex-bonded curled hair in sheet form, (*b*) foamed rubber, a soft, expanded rubber of cellular form, (*c*) foamed plastic, based on polyurethane materials, of open or closed cell structure.

When a package suffers an impact after a free fall, the outer container is brought to rest suddenly, but the kinetic energy possessed by the inner one is absorbed by the surrounding cushions, which therefore apply an increasing force as they are compressed and cause a suitable deceleration of the inner article to bring it to rest. The peak deceleration (peak G) occurs at the point of maximum compression, and this determines the peak stress in all the components, and hence the possibility of fracture. From the point of view of economy of size and material, the cushion design should be such that this point occurs at a reasonable compression (about 70–80%) but it must avoid the region where the cushion 'bottoms' as the stiffness suddenly rises as it becomes fully compressed. The total force needed to compress any resilient material is not only that due to the static stiffness, but also any component arising from energy losses or damping. The dynamic stiffness characteristic therefore lies above the static curve during the forward stroke, and below it during the return, additional energy being needed to compress and less being returned. Many factors can contribute to the damping losses, but whatever they are, the resultant is almost certain to be frequency dependent. In addition, materials of a closed-cell nature exhibit a component of stiffness due to the pneumatic effect of compressing the gas in the cells, and this again will differ between slow and fast compressions. Dynamic stiffness is therefore the characteristic which controls the peak deceleration experienced by a given article when dropped from a certain height and it is essential to measure this under conditions closely related to those of the actual impact.

2. Determination of dynamic response

In view of the lack of knowledge of the dynamic performance of the cushioning materials, packaging design has largely been based on static curves, which usually give errors on the safe side as the additional dynamic effects all contribute extra energy losses, and thus lead to lower figures of peak G. Packaged systems in general do not readily lend themselves to theoretical analysis owing to the presence of the non-linear stiffness, but one method which has shown promising results is treatment via an electronic analogue computer (Venning and Guttridge 1960, private communication).† This replaces the original mechanical system by an analogous one of electrical voltages, and non-linear stiffness terms are easily simulated if the form is known. If the static stiffness curve is used, the solution of the problem is correct only for that particular law, so that the dynamic stiffness should be used, but this is itself dependent to a large extent on the solution, as the energy losses causing the dynamic effects are frequency dependent. For this reason, an analogue system was devised incorporating a unit utilizing a sample of cushioning material to provide a 'real' stiffness element, which therefore not only helped in the simulation of the correct system to obtain solutions of acceleration–time, etc., but also gave the dynamic compression curve of force–displacement of the cushion appropriate to the solution.

The response of the simple system of Fig. 1(*a*) can be determined using a number of analogue units connected as in Fig. 1(*b*), with the static stiffness law simulated by unit D connected between P and Q. The mass m_1, travelling with an initial velocity $\dot{x}_0$ is brought to rest by a cushion which has

* Now with the Department of Physical Chemistry, University of Cambridge.

† Proceedings of Symposium at Royal Radar Establishment, Malvern, April 1960, circulated by the Royal Radar Establishment.

a static stiffness law $F(x)$, the motion being governed by the differential equation $m_1\ddot{x} + F(x) = 0$ or $\ddot{x} = -F(x)/m_1$. In the analogue system, voltages $V_{\ddot{x}}$, $V_{\dot{x}}$ and V_x are made proportional to acceleration $\ddot{x}$, velocity $\dot{x}$ and displacement x. The functions of the units are arranged so that unit A introduces the coefficient m_1, units B and C integrate with respect to time, and unit D transfers its input to output via any desired static law. The closed loop connection is such that unit A makes $V_{\ddot{x}} = -V_{F(x)/m_1}$ which is the analogue form of the differential equation of the system, and outputs are therefore obtained proportional to $\ddot{x}$, $\dot{x}$ and x.

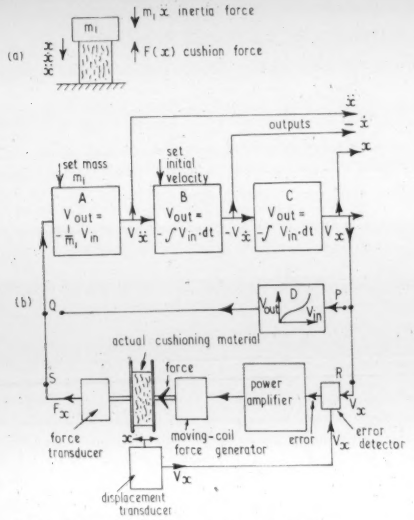


Fig. 1. Analogue computer solution of the dynamic system.

To incorporate the real stiffness element, unit D is replaced by a dynamic unit, shown between points R and S. The requirement for this is that at any instant within the compression cycle when the displacement is x_1, it compresses a sample of the cushion material to that displacement and produces an output proportional to the force used, i.e. the $V_{F(x)}$ is now obtained via the actual cushion instead of using a static curve. The compression force is obtained from a moving-coil generator, driven by a power amplifier, and a servo system ensures that the correct displacement is obtained at each instant. Details of the various parts are as follows:

2.1. *Moving-coil force generator*

This is based on a Goodman's Vibration Generator, Model 390A, with the coil and magnet system modified to give a travel of 2 in. The pole piece was modified to maintain a gap length of 0·5 in. and provided with three nylon inserts which rub lightly against the inside of the coil former and locate it within the gap. The thrust is then coupled from the coil to the pressure plate via a rod supported in a plain brass-bush bearing. (A ball-bearing bushing at this point gave a large amount of 'noisy' vibration on the shaft.) The coil is wound with a single layer of 22 gauge enamelled-copper wire, with 17 active turns in the magnetic field, out of a total of 85 turns, giving a total resistance of 1·0 ohm (with flexible leads). The magnetic field strength in the gap is approximately 1·1 Weber/m^2 (11 000 gauss), and the force law is 0·85 lb/amp. of coil current. The pressure plate is made of magnesium to keep the weight to a minimum, and the total moving mass is 7 oz.

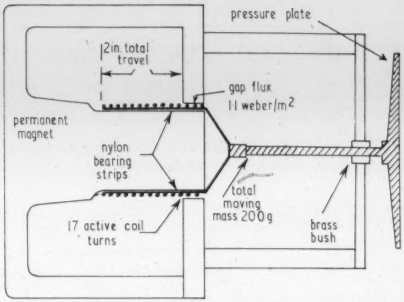


Fig. 2. Cross section of moving-coil force generator.

2.2. *Force transducer* (Fig. 3)

The force transmitted through the cushion is measured by a variable-capacitance pick-up having a high natural frequency (above 5 kc/s), so that it operates as a stiffness-controlled device with the displacement of the moving plate proportional to applied force. The pick-up is used in conjunction with a frequency-modulated system where change of the pick-up capacitance alters the frequency of a Franklin oscillator, the carrier frequency being 490 kc/s. A buffer stage is followed by an f.m. detector utilizing an EQ80 nonode valve, the output of this being passed to a d.c. amplifier. The variable capacitance pick-up has a nominal capacitance of 80 pF, and an applied force of 4 lb causes a variation of 1 kc/s in the oscillator frequency, with a resultant output of 5 v at maximum gain. Short term drift at the output is of the order of 30 mv, due chiefly to the d.c. amplifiers. It is essential to measure the force applied to the cushion either at the pressure plate or at the back plate, as it differs from the total force generated at the coil by the amount used to accelerate the moving mass, only becoming equal to it when the latter is at rest. The natural frequency of the back-plate and transducer is kept high by securing the lower edge of the plate to the main frame, so that it behaves as a vertical cantilever.

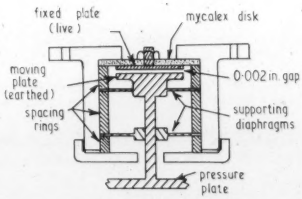


Fig. 3. Force transducer.

2.3. *Displacement transducer*

This is also a variable-capacitance pick-up, consisting of a rod carried by the pressure plate, moving parallel to a fixed insulated plate, at a separation of 0·075 in. Although a relatively crude device, the linearity of output over a travel of 2 in. was quite sufficient. The carrier frequency was 530 kc/s, and the f.m. circuits were identical to those used for the force transducer. The chief disadvantage was the presence of an out-of-balance mass at the top of the pressure-plate, together with the need for an additional guide, and the rod-and-plate design has now been replaced by a pair of co-axial cylinders mounted about the push-rod on the main centre-line, between the coil and the main bearing.

2.4. *Power amplifier*

As this is within the servo loop, a linear power amplifier is required, the general requirements being:

(i) D.C. coupling, as the low-frequency components of pulses in the range of 100 ms would suffer too much phase-shift in any a.c. coupled amplifier.

(ii) Current output in the range of amperes.

(iii) Load resistance of 1·0 ohm, with a variation of impedance during movement of the coil due to the back e.m.f. generated, so that the output impedance of the amplifier must be high, acting as a constant-current source.

(iv) No d.c. component of current in the load under quiescent conditions.

Transformer-coupled circuits are unfortunately ruled out by the need for a d.c. response, so the design employed a transistor amplifier, using parallel connections of p–n–p germanium power transistors for the output stage. Various push–pull, bridge and cascade circuits were considered from the points of view of current handling capacity and efficiency, and from the need to obtain a high output impedance. The most suitable arrangement eventually used two power transistors (Newmarket Type V60/30P) in parallel as a grounded-emitter driver stage, feeding a further ten transistors connected in grounded-collector as the main output stage to supply the 1·0 ohm load. Backing-off current was fed through the load to balance the quiescent current and a separate base bias supply kept collector current independent of changes in load impedance.

Working from a supply voltage of ±12 v, overall performance is

Current gain = 500.
Peak–peak output in Class A, 10 a.
Zero-peak output approximately in Class B, with uni-directional pulse 15 a.
Bandwidth, d.c., 2 kc/s.

2.5. *Servo loop*

Velocity feedback was added to improve the performance, the velocity term being obtained by differentiating the displacement output. With an optimum setting of this, pulses of 100 ms were reproduced with little distortion and overshoot, but it was considered undesirable to use shorter pulses than this as the system resonated at 45 c/s.

3. Typical experimental results

Tracings of oscillographs of some typical results for samples of various cushioning materials are shown in Fig. 4. In each case, the top trace is that of compression force against time, the middle trace is of displacement against time, and the lower one is that of compression force against displacement.

As a test of performance, a small linear metallic spring was used, giving the results shown in Fig. 4(*a*). The stiffness

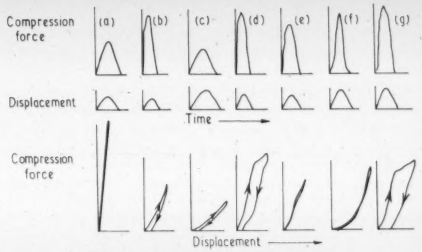


Fig. 4. Typical experimental results.

measured on a static load test was 18·5 lb/in. and that obtained from the force–displacement curve is 18·6 lb/in. for a pulse length of 168 ms. The chief errors of measurement arise in making the final measurements from the photographed traces. The narrow hysteresis loop obtained in this case represents the small energy loss arising in the spring.

Cushion details for other samples are given in the Table

(1)	(2)	(3)	(4)	(5)	(6)	(7)
4(*b*)	Rubberized hair 4 lb/ft^3	2·0×3·3	1·8	125	0·42	3·8
4(*c*)	Rubberized hair 4 lb/ft^3	2·0×3·3	1·8	214	0·69	6·8
4(*d*)	Foamed polyurethane (polyester) 2 lb/ft^3 open cell	2·8×1·7	1·7	136	0·83	9·15
4(*e*)	Foamed P.V.C. 12 lb/ft^3 open cell	2·0×2·0	1·0	145	0·83	9·73
4(*f*)	Fibroceta cellulose wadding	2·6×3·1	0·9	152	0·83	11·4
4(*g*)	As in (*d*), enclosed in sealed P.V.C. envelope	2·8×2·0	1·4	163	0·83	15·6

1, Figure; 2, Material; 3, Area (in^2); 4, Thickness (in.); 5, Force pulse length (ms); 6, Peak compression (in.); 7, Peak force (lb).

4. Comments

The results shown above demonstrate the possibilities that exist in this method of measuring the dynamic performance of the cushion samples. The instrumentation for force and displacement transducers is satisfactory, although the displacement transducer has since been replaced by electrodes consisting of two concentric cylinders mounted on the main axis between the coil and the forward bearing. The chief limitation at present is that of available force, which arises from the need to use a linear power amplifier in association with the computing units. For testing with certain force-pulse waveforms, such as square, or half-sine, then the linear amplifier can be dispensed with and a switching and shaping

arrangement used to derive current pulses of much greater magnitude. This is being pursued, with the object of using larger cushion samples, in area and thickness, with compressions to bottoming point. Measurement accuracy from the final traces is poor, as in all photographic work, but readings of the peak values could be improved with peak reading meters.

The force waveform shows a definite higher-frequency oscillation on some traces (e.g. Fig. 4(*e*)), which would appear to be due to compression waves in the bulk material, rather than to spurious responses of the system itself. The effect of compression waves in this method of testing differs from that experienced in actual cushioning use as they are recorded after a single transmission through the cushion thickness, instead of being reflected back to the compression plate again. The frequency response of the present system is not considered high enough to investigate these with any degree of accuracy.

The waveforms of Fig. 4(*b–g*) show clearly the force–time responses of different materials, together with the corresponding force–displacement curves. All exhibit non-linear stiffness characteristics, with a definite area enclosed within the force–displacement curve, showing the degree of energy dissipation during a single half-cycle.

Acknowledgments

The work was carried out in the Department of Electronics, University of Southampton, with the financial assistance of Messrs. Wilmot, Mansour and Co. Ltd., Southampton. The authors would like to acknowledge the interest of Professor E. E. Zepler, and of Mr. S. C. Schuler, Royal Radar Establishment, Malvern, and Mr. Alexander Orba.

References

SCHULER, S. C., 1958, *J. Instn Elect. Engrs*, **4** (38) 76.

New Books

Nondestructive Testing. By W. J. MCGONNAGLE. (New York, Toronto, London: McGraw-Hill, 1961.) Pp. xi + 455. Price £5 16*s*. 6*d*.

In the preface the author observes that education in the field of nondestructive testing is a primary requirement for technology and that, as yet, few institutions of higher learning offer formal courses in this field. He lives in the United States of America but it prompts the observation that consideration might be given in this country to the provision of courses in the subject. For example, nondestructive testing could suitably form part of the Higher National Certificate courses in Applied Physics or Engineering at some technical colleges. It could be included in certain technical degree courses. At the end of the book the author rightly warns against some of the pitfalls of an unimaginative and uninformed approach to nondestructive testing—for example "The idea is too prevalent that all one has to do is to get an instrument and one's troubles are over". He states that his aim has been to discuss basic principles, correct application, technique, and the results to be expected and interpreted.

Between this 'beginning' and 'end' of the subject we find a very wide coverage of many established physical methods for the examination of materials. The important methods covered are classifiable under Visual testing, Pressure and leak testing, Liquid penetrant inspection, Thermal methods, Radiography (x-ray and gamma ray), Ultrasonics, Dynamic testing, Magnetic, electrical and eddy currents. 'Other useful techniques' and a separate chapter on thickness measurements complete the survey. As regards the main topics the treatment is very thorough. The numerous illustrations are well chosen—some of them would give the technologist useful knowledge about an instrument whilst others could well be used as a guide to interpretation. In this respect the book is practical without being a laboratory bench manual. The treatment of basic physical principles is adequate although one might sometimes have hoped for a little more (a real difficulty in a book which has such extensive terms of reference). The reviewer recognizes that anyone coming to use the book will already have some basic knowledge of physics and will be able to see how the technology is built on these foundations.

It follows that the author has provided a sound and wide foundation for nondestructive testing regarded as a branch of applied physics or even as a technology in its own right. The price? This could be a deterrent but not to seriously involved students and certainly not to libraries, individual practitioners, or teachers, and there will or should be some more of the latter.

K. W. ANDREWS

Simple viscometer for use with low melting point metals

by D. M. COTTINGHAM,* B.Sc., The General Electric Company Limited, Central Research Laboratories, Hirst Research Centre, Wembley, Middlesex

MS. received 12*th May* 1961

Abstract

The text describes a viscometer suitable for making relative measurements over a wide range of viscosities and over a temperature range of 20° *to* 600° C. *The viscometer is calibrated with standard liquids and has been designed particularly for low melting point metals. The operating techniques and sources of error are discussed. An empirical equation relating viscosity to the variable parameters is given and results of experiments on lead are included.*

Introduction

Several problems must be overcome before satisfactory measurements can be made of the viscosities of liquid metals. The viscosity is usually of the order of a few centipoises and thus a fairly sensitive piece of apparatus must be used. Very large errors often arise from solid oxide film: these are generally most troublesome in apparatus where the viscosity is measured either by the damping of an oscillating pendulum or in apparatus of the Couette type. When the flow through a capillary tube is used to determine the viscosity there is the disadvantage that oxide will block or partially block the capillary tube.

The apparatus described in this paper was designed primarily to overcome the problems associated with oxide films. The variable geometry of the apparatus allows a wide range of viscosities to be measured. Although organic liquids were used only for calibration of the apparatus it would appear to be suitable for measurements on a wide variety of these liquids.

Principle

The apparatus (British Patent Application No. 22669/59) shown diagrammatically in Fig. 1 consists essentially of a drum which may be rotated inside a metal tank. The flat end-faces of the drum are in light contact with the sides of the tank and the bottom of the drum is separated by only a small clearance from the bottom of the tank. Thus the arrangement makes effectively two compartments joined by a narrow duct below the bottom of the drum.

When the drum is rotated about its horizontal axis, liquid is dragged through the narrow duct and a head of liquid builds up in one compartment: this head forces some of the liquid to flow back through the bottom half of the duct. A scraper lightly pressed against the top of the drum prevents liquid being dragged across from one compartment to the other. A measure of the viscosity of the liquid is given by the equilibrium head of liquid produced when equal volumes per unit time are flowing in opposite directions through the duct.

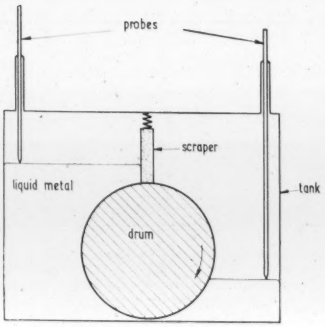


Fig. 1. Diagram illustrating the principle of the viscometer.

Construction of the apparatus

Mild steel was used for all the major components of the apparatus since it is corrosion resistant to bismuth, one of the metals being studied, at temperatures up to about 500° C.

Conventional lubrication of the interface between the end faces of the drum and the tank was not possible and to reduce friction the drum was faced with graphite. Graphite is compatible with bismuth but the coefficient of expansion is much less than that of mild steel. Thus as the temperature of the apparatus increased a small gap developed between the sides of the tank and the graphite end-faces of the drum. To compensate for the expansion of the tank the scraper was constructed with a tenon joint which was expanded by means of a blade spring. The springs within the tank were made from either tungsten sheet or wire to reduce corrosion and to ensure good creep properties. The graphite scraper was lapped in to follow the contour of the drum by running the apparatus for a few hours at the desired temperature.

Since it was essential that the apparatus should not distort asymmetrically when heated, all the steel used for construction was first skinned and then fully annealed before final machining and grinding. The apparatus was made sufficiently massive to reduce creep to a minimum and to increase heat conduction round the tank thus reducing variations in temperature to a minimum.

Figure 2 shows a cross section of the apparatus. From this it will be seen that the drum is supported on a shaft which rotates in ball-race bearings. These ball races are set eccentrically in the bearing housings and the latter may be rotated in the end pillars of the apparatus. Because the shaft is eccentric relative to the bearing housing, rotation of the latter causes the shaft to be either raised or lowered relative to the bottom of the tank. The drum is keyed to

* Now at Tube Investments, Research Laboratories, Hinxton Hall, Cambridge.

the shaft which is driven by an electric motor through two reduction gears. Above the drum is the graphite scraper and it is held in contact with the drum by means of the two

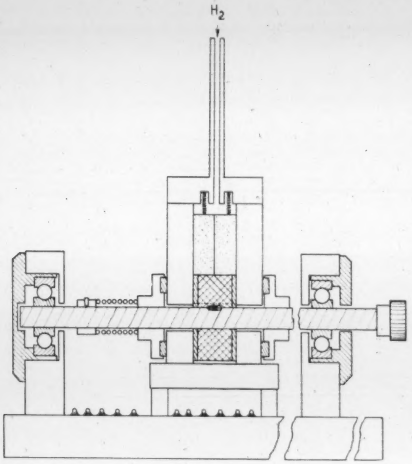


Fig. 2. Diagrammatic section of the rotating drum viscometer. Scale: 1 cm ≡ 1·25 in.

compression springs. On either side of the tank are shown liquid-metal seals which prevent liquid metal leaking out of the tank if the graphite drum faces become worn.

The tank is heated by a 2·3 kw tubular strip element fitted to the sides of the tank, and a thin element similar to the type used in a domestic iron is clamped to the bottom of the tank.

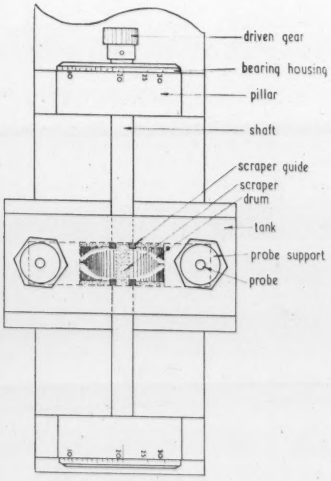


Fig. 3. Diagrammatic representation of the rotating drum viscometer. Scale: 1 cm ≡ 1·25 in.

The top face of the tank is thermally lagged with insulating material.

The levels of the liquid metal within the tank are measured by probes the positions of which are indicated in Figs 1 and 3. Each probe consists of a tungsten rod which is brazed into a steel rod. The latter passes through a neoprene O-ring gas seal in the probe support. The probes are thus electrically insulated from the tank. The level of the liquid metal is determined by a simple electrical circuit which is described in the next section.

The base of the apparatus and upper portions of the probe supports are water-cooled and the shaft is air-cooled at its ends.

Method of operation

The desired gap between the drum and base of the tank is set by rotating the bearing housings. Angular displacements corresponding to changes of gap of 0·001 in. are marked on the bearing housings. The probes are lowered slowly until their pointed ends just contact the surface of the liquid metal. This completes an electrical circuit and an indicator bulb lights for each probe. Cursors which are attached to the upper ends of the probes are set to zero on their accompanying measuring scales. The electric motor is now switched on and the speed of rotation of the drum is adjusted to the slowest of six set speeds. The probes are next moved until their tips are again just in contact with the metal surface. After a few minutes a check is made to see that the readings indicated are constant and that the liquid metal has reached its equilibrium height. The displacements of the probes from zero are measured and added together to give the head of liquid metal corresponding to a given gap, speed of rotation, and temperature. The probe measurements are repeated for the other five speeds and then the gap is altered and the cycle of measurements begins once again. For gaps in excess of 0·005 in. each cycle requires about fifteen minutes. Smaller gaps may require up to several hours to complete the cycle of measurements as some time is required before the metal reaches its equilibrium head for a given speed.

Measuring devices and errors

The speed of the motor and hence the rotational speed of the drum is measured by a stroboscopic disk illuminated at mains frequency. The motor speed is controlled electronically and is constant to within 0·1% of the set speed. Fluctuations in the mains frequency can be as much as 1% and so errors in the rotational speed of the drum are of about the same value. The head of liquid can be measured to an accuracy of 0·1 mm and one of the largest sources of error comes from this measurement; but these errors are averaged out on the graph of head of liquid plotted against speed and a specimen graph is given in Fig. 4.

Although the apparatus is designed to measure viscosities between one and more than a thousand centipoises it will also measure accurately very small changes in viscosity. The final derived viscosity is a graphical solution of about 50 readings using eight or more duct sizes.

The temperature of the liquid metal is measured by a thermocouple in the centre of the bottom plate of the tank about 0·01 in. from its upper surface. At the melting point of bismuth the thermocouple recorded 271° c and thus it would appear that the accuracy of the temperature measurements is within 2° c.

One possible source of error is the decrease in gap between drum and tank caused by the thermal expansion of various

parts of the apparatus. The gap heights marked on the bearing housings were determined by means of a dial gauge and feeler gauges with the apparatus at room temperature. At 450° C the zero error is about −0·009 in., but this error can be estimated fairly accurately by analysis of the results obtained with various gap widths.

that of the standard liquid. An example of these plots is shown in Fig. 5 where the results for lead at three different temperatures are given.

Acetone was used to standardize the apparatus and the viscosity of methyl alcohol was measured relative to the known value for acetone. The results are in close agreement

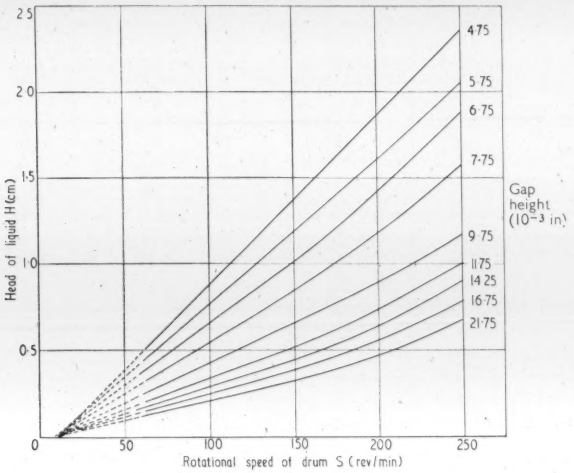


Fig. 4. Results for lead, 393° C.

Results

The empirical expression relating viscosity η to the difference in head H of fluid between the two chambers, the height of the gap clearance D, the density of the fluid ρ and the speed of rotation of the drum S is

$$H = \frac{K_1 \eta S}{\rho D} + K_2. \qquad (1)$$

This expression holds for wide variations in operating conditions but its limitations are discussed in the next section. K_2 is found to vary with the gap height, but it is approximately zero for large gap heights.

For any particular value of D the head of liquid H is plotted against the speed of rotation of the drum S as in Fig. 4.

The slope of each line for a particular gap is given by

$$\frac{dH}{dS} = \frac{K_1 \eta}{\rho D}. \qquad (2)$$

By plotting dS/dH against D a straight line is obtained of gradient $G = \rho/K_1\eta$.

Using subscripts 's' for the standard liquid and 'm' for metal whose viscosity is to be measured,

$$\eta_m = \frac{\rho_m}{G_m}\frac{G_s}{\rho_s}\eta_s. \qquad (3)$$

Thus the viscosity of a liquid may be determined by comparing the slopes of the plot of dS/dH against D with those published in *International Critical Tables* as shown in the table below.

The apparatus was further tested with pure lead and bismuth. The results for these metals are given in the table and for lead they agree very closely with the published data. Previously reported values for bismuth show a wide scatter and our results are about 70% above those of Sauerwald and Tropler (1926), but the slope of the graph of viscosity against temperature is approximately the same as that given by these authors.

Table showing measured values of viscosity of some liquids and values from International Critical Tables

Liquid	*Temperature* (° C)	*N (measured)* (cP)	*N (tables)* (cP)
Acetone	20	Standard	0·325
Methanol	20	0·567	0·576
Bismuth	368	2·3	1·47
Bismuth	404	2·25	1·35
Bismuth	440	2·06	1·30
Lead	342	2·56	2·575
Lead	393	2·36	2·37
Lead	440	2·21	2·15

Discussion

For Eqn (1) to be satisfied the conditions of the experiment must be such that the flow through the duct is laminar. With a gap of 0·020 in. tests suggest that some transition from laminar flow to turbulent flow occurs with lead at a drum speed of 150 rev/min, and for these conditions the Reynolds

number defined as DSp/η is 1300. This breakdown is illustrated in Fig. 4 where it may be noted that the rotational speed of the drum for transition to occur increases as the gap decreases. At least the first four points on the curve for any gap height lie on straight lines, which indicate laminar flow.

It may be noted that the extrapolated line for any given gap height does not cut the axis of 'rotational speed of drum' at zero and the intercept gives the term K_2 in Eqn (1). Furthermore as the gap decreases the K_2 increases. A possible reason for this shift is that the scraper does not quite prevent all the liquid carried through the duct from returning over the top of the drum. For large gaps relatively large volumes of liquid are passing in both directions through the duct below the bottom of the drum, and thus the error due to the small amount returning between the scraper and top of the drum is small. But this error increases as the gap height decreases and Fig. 4 shows that with the present apparatus for gaps below 0·004 in. the error would be unacceptable. In the measurements on lead only the results for gap heights of 0·005 in. upwards have been used and these are plotted in Fig. 5. The points fall very nearly on straight lines and illustrate the reasonable accuracy obtained from the instrument. The graphs show the effect of temperature on the viscosity of lead as measured by the viscometer.

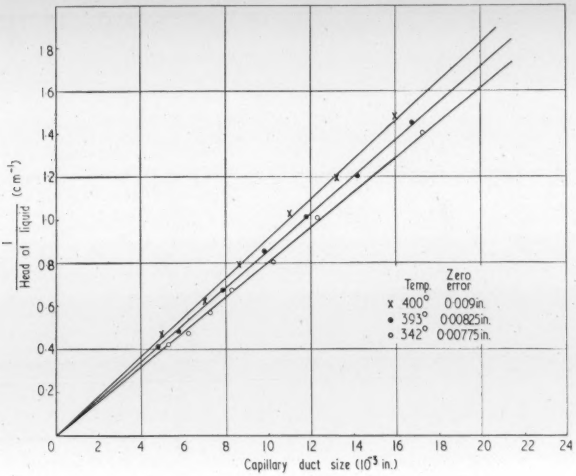


Fig. 5. Variation of viscosity of lead with temperature.

Acknowledgment

This work was supported by The United Kingdom Atomic Energy Authority, Harwell, and the author wishes to thank the Director for permission to publish the paper.

Reference

SAUERWALD, F., and TOPLER, K., 1926, *Z. Anorg. Chem.*, **157**, 117.

Some electromechanical effects on dielectrics

by K. C. KAO, M.Sc., Ph.D., A.M.I.E.E., Brush Electrical Engineering Company Limited, Loughborough, Leics.

MS. received 7th April 1961, in revised form 9th August 1961

Abstract

On the basis of the principle of minimum potential energy, mechanical forces in electrically stressed dielectrics have been derived. These forces are appreciable in strong electric fields. A brief discussion is given with particular reference to their effects on the dielectric breakdown.

1. Introduction

MECHANICAL forces in electrically stressed dielectrics have so far received very little attention. The purpose of this paper is to show that these forces, which tend to reduce the potential energy in an electrically stressed system to a minimum, are appreciable in strong fields and that their effects on dielectrics are of some importance.

2. Mechanical forces and their effects

In an electrostatic system the forces of electric fields tend to increase the capacitance in order to reduce the potential difference between charges irrespective of whether constant charges or constant voltages are applied. This is the principle on which the derivation of mechanical forces in electrically stressed dielectrics is based. For mathematical simplicity it is assumed that the dielectrics under consideration are isotropic and linear, that the initial charges in the system are fixed, and that the mechanical work is done at the expense of the stored energy in the field. Heat losses and other effects such as thermal agitation and gravitation are ignored. Thus the change of stored energy U in the electric field due to the introduction of a dielectric body of volume v and permittivity ϵ_2 into a dielectric medium of permittivity ϵ_1 can be calculated by the following equation (Stratton 1941):

$$U = \tfrac{1}{2}\int_v (\epsilon_2 - \epsilon_1)E_1 E_2 dv \qquad (1)$$

where E_1 is the field strength in the region of the dielectric medium where the body is subsequently introduced and E_2 is that in the body itself. The stored energy in the field is decreased when U is positive, and increased when it is negative. Using this equation, mechanical forces are derived and their effects described together with quantitative indications of their significance.

2.1. *The forces acting on the boundary between two dielectrics*

At the boundary between two dielectrics the electric flux is usually refracted, and the tangential components of the field strengths in dielectrics 1 and 2, E_{1t} and E_{2t}, as well as the normal components, E_{1n} and E_{2n}, follow the relationship

$$E_{1t} = E_{2t} \qquad (2)$$

$$\epsilon_1 E_{1n} = \epsilon_2 E_{2n} \qquad (3)$$

where ϵ_1 and ϵ_2 are the permittivities of dielectrics 1 and 2 respectively.

From Eqns (1) and (2) the decrease in stored energy due to the expansion of a small volume dv of dielectric 2 into dielectric 1 for the tangential component of the field is

$$dU_t = \tfrac{1}{2}(\epsilon_2 - \epsilon_1)E_{1t}^2 dv. \qquad (4)$$

Thus the force acting on the boundary is

$$dF_t = \nabla(dU_t) = \tfrac{1}{2}(\epsilon_2 - \epsilon_1)\nabla E_{1t}^2 dv. \qquad (5)$$

Since $dF_t/dv = \nabla P_t$, the mechanical force per unit area or mechanical stress due to the tangential component of the field is

$$P_t = \tfrac{1}{2}(\epsilon_2 - \epsilon_1)E_{1t}^2. \qquad (6)$$

Similarly it can be shown that the mechanical stress due to the normal component of the field is

$$P_n = \tfrac{1}{2}\frac{\epsilon_1(\epsilon_2 - \epsilon_1)}{\epsilon_2}E_{1n}^2. \qquad (7)$$

These mechanical stresses, whose directions are normal to the boundary, tend to cause the dielectric having higher permittivity to move into the space occupied by the other.

Most dielectric materials used in insulation systems have relative permittivities ranging from 1 to 8. Some typical magnitudes of these mechanical stresses in practical cases are given in Table 1. Liquid or solid dielectrics are generally able to withstand intrinsically an electric stress of several MV/cm for a short period of microsecond order, but may not be able to resist a mechanical impact of about 10 times atmospheric pressure which may easily occur, as indicated in Table 1, when subjected to transient overvoltages such as lightning or switching surges.

Table 1. *Some typical magnitudes of mechanical stresses*

Permittivity ϵ_1/ϵ_0	ϵ_2/ϵ_0	*Tangential field* E_{1t}(MV/cm)	*Stress* P_t(lb/in^2)	*Normal field* E_{1n}(MV/cm)	*Stress* P_n(lb/in^2)
4	8	1	25·7	1	12·8
4	8	2	102·8	2	51·2
2	8	1	38·4	1	9·6
2	8	2	153·6	2	38·4

ϵ_0 is the permittivity of free space.

2.2. *The force acting on conductor surfaces*

Since the electric flux is normal to the conductor surfaces, there exists a Coulombic force tending to push the conductors with charges of opposite polarity together. According to the principle of minimum potential energy the mechanical stress P_e acting on the conductor surfaces is equal to the energy density stored in the field. Thus

$$P_e = \tfrac{1}{2}\epsilon E^2 \qquad (8)$$

where E is the field strength and ϵ is the permittivity of the dielectric.

This mechanical stress will result in a compression of the dielectric and its effect on the breakdown strength of solid dielectrics has been reported by many investigators (Cooper and Wallace 1953, Stark and Garton 1955). For example, a dielectric such as a plastic material having relative permittivity of 4 stressed at a field of 2 MV/cm may suffer a compressive pressure of 102·8 lb/in^2. This pressure, which increases proportionally with permittivity and the square of the field strength, may be sufficient, especially under transient overvoltage conditions, to cause reduction of thickness or fracture of the insulation, followed by dielectric breakdown.

2.3. *The force elongating a bubble or a globule*

A small bubble of gaseous or vapour phase, or a globule of liquid phase in a fluid dielectric always takes on a spherical shape due to surface tension. As the field strength inside a spherical bubble (or globule) is

$$E_2 = \frac{3\epsilon_1}{2\epsilon_1 + \epsilon_2} E_1 \tag{9}$$

the change in stored energy due to the presence of a spherical bubble of permittivity ϵ_2 into a fluid dielectric of permittivity ϵ_1, from Eqn (1), is

$$U_s = \frac{\epsilon_1 {E_1}^2}{2}\left\{\frac{3(\epsilon_2 - \epsilon_1)}{2\epsilon_1 + \epsilon_2}\right\}\left(\frac{4}{3}\pi r^3\right) \tag{10}$$

where r is the radius of the bubble and E_1 is the field strength in the region of the fluid dielectric which the bubble subsequently occupies. It is assumed that a spheroidal shape is a good approximation to that into which a spherical bubble may be expected to deform, and that the volume of the bubble remains unchanged whilst its shape deforms from a sphere into a spheroid. As the field strength inside a spheroidal bubble (Bottcher 1952) is

$$E_2 = \frac{\epsilon_1}{\epsilon_1 - (\epsilon_1 - \epsilon_2)G} E_1 \tag{11}$$

the change in stored energy due to the presence of a spheroidal bubble, from Eqn (1), is

$$U_e = \frac{\epsilon_1 {E_1}^2}{2}\left\{\frac{\epsilon_2 - \epsilon_1}{\epsilon_1 - (\epsilon_1 - \epsilon_2)G}\right\}\left(\frac{4}{3}\pi r^3\right) \tag{12}$$

where

$$G = \frac{ab^2}{2}\int_0^\infty \frac{ds}{(s + a^2)^{3/2}(s + b^2)}, \tag{13}$$

$2a$ and $2b$ are the lengths of the principal axes of the spheroidal bubble, axis $2a$ having the direction of the electric field, and s is the parameter of a family of spheroids all confocal with the bubble surface $s = 0$.

From Eqns (10) and (12) the work done at the expense of the stored energy for the bubble deformation is

$$U = U_e - U_s = \frac{\epsilon_1 {E_1}^2}{2}\frac{(\epsilon_2 - \epsilon_1)^2(1 - 3G)}{(2\epsilon_1 + \epsilon_2)\{\epsilon_1 - (\epsilon_1 - \epsilon_2)G\}}\left(\frac{4}{3}\pi r^3\right). \tag{14}$$

It is obvious that for $U > 0$, G is less than $\frac{1}{3}$, which corresponds to a prolate spheroid, that is, a spherical bubble will elongate in the direction of the field. To calculate the force for bubble elongation, it is necessary to know the decrease in stored energy when a sphere is only slightly deformed by elongation of the major axis from the original $2r$ to $2(r + da)$. Since

$$r^3 = ab^2 \tag{15}$$

$$a = r + da. \tag{16}$$

Hence

$$\frac{b^2}{a^2} = 1 - \frac{3da}{r} \tag{17}$$

the terms of higher powers in the binomial expansion being ignored. Integration of Eqn (13) and substitution of Eqn (17) into it give

$$G = \frac{1}{3} - \frac{da}{r}. \tag{18}$$

Substitution of Eqn (18) into Eqn (14) gives the decrease in stored energy due to a bubble with its major axis extended from $2r$ to $2(r + da)$

$$dU = 6\pi r^2 \epsilon_1 \left(\frac{\epsilon_2 - \epsilon_1}{2\epsilon_1 + \epsilon_2}\right)^2 {E_1}^2 da. \tag{19}$$

Thus the elongation force is

$$F_e = \frac{dU}{da} = 6\pi r^2 \epsilon_1 \left(\frac{\epsilon_2 - \epsilon_1}{2\epsilon_1 + \epsilon_2}\right)^2 {E_1}^2. \tag{20}$$

The elongation force depends on the size of the bubble irrespective of whether the permittivity of the bubble is larger or smaller than that of the fluid. For example an air bubble of radius 1 micron formed in transformer oil whose relative permittivity is about 2 will suffer an elongation force per unit area of about 13·6 lb/in^2 at a field of 1 MV/cm. This force is sufficient to overcome the surface tension of the bubble and to cause bubble elongation. Kao and Higham (1961) have suggested that the breakdown of liquids commences in the bubble when its elongated axis and voltage across it are suitable. It was observed in the author's laboratory that in transformer oil the field needed to elongate an artificial air bubble is much higher than that needed to elongate a water globule, as expected from Eqn (20).

2.4. *The dielectrophoretic force*

Dielectrophoresis is defined as the motion of particles in a non-uniform field depending only on the forces produced by the polarized particles. The dielectrophoretic force tends to draw the particles whose permittivities are higher than that of the dielectric medium from the weak to the intense field, and eject those of lower permittivities from the intense to the weak field. On the assumption that the particles are spherical in shape, the change in stored energy due to the presence of a spherical particle of radius r and permittivity ϵ_2 in a dielectric medium of permittivity ϵ_1, stressed at a field E_1, can be calculated by Eqn (10). Thus the dielectrophoretic force is

$$F_d = 2\pi r^3 \epsilon_1 \left(\frac{\epsilon_2 - \epsilon_1}{2\epsilon_1 + \epsilon_2}\right) \nabla {E_1}^2. \tag{21}$$

Consider a spherical electrode of radius R, having a potential V_1, with respect to zero potential at a concentric sphere of infinite radius. The field strength or voltage gradient at a distance h from the centre of the electrode is

$$E_1 = \frac{RV_1}{h^2}. \tag{22}$$

Thus

$$F_d = \frac{8\pi r^3 \epsilon_1 R^2 V_1^2}{h^5}\left(\frac{\epsilon_2 - \epsilon_1}{2\epsilon_1 + \epsilon_2}\right). \quad (23)$$

This force will be opposed by a viscous drag which, according to the Stokes law, is $6\pi\eta r v$, where η is the viscosity of the medium and v is the velocity of the particle. Hence the velocity of the particle in such a field is

$$v = \frac{4r^2\epsilon_1 R^2 V_1^2}{3\eta h^5}\left(\frac{\epsilon_2 - \epsilon_1}{2\epsilon_1 + \epsilon_2}\right) \quad (24)$$

and the time required for the particle to travel from a distance h to the electrode surface is

$$t = \frac{\eta(h^6 - R^6)}{8r^2\epsilon_1 R^2 V_1^2}\left(\frac{2\epsilon_1 + \epsilon_2}{\epsilon_2 - \epsilon_1}\right). \quad (25)$$

For a spherical electrode of radius 2 cm immersed in transformer oil whose viscosity is 37 centistokes, some typical travelling times of particles under various conditions are given in Table 2. It can be seen that some particles may take less than 1 ms to travel 1 cm, but some may take more than 12 minutes depending on the field strength and the size of the particles. This indicates that the severity of the dielectrophoretic effect depends on the duration of the applied field. As impurities migrated to the intense field may initiate ionization or cause bubble formation, it should be expected that the breakdown strength of transformer oil under impulse voltage conditions is higher than that under continuously stressed conditions. The dielectrophoretic effect may, on the other hand, be used to produce pumping action of non-conducting liquids, to cause separation of the components in suspension in a fluid dielectric, to cause precipitation, or to produce mixing (Pohl 1958).

Table 2. *Some typical travelling times of polarized particles*

ϵ_1/ϵ_0	ϵ_2/ϵ_0	r (mm)	h (cm)	t (s) for $V_1 = 100$ kv	t (s) for $V_1 = 1000$kv
			3	745·00	7·45
		0·01	4	4510·00	45·10
			5	17400·00	174·00
			3	7·45	0·0745
2	8	0·10	4	45·10	0·451
			5	174·00	1·74
			3	0·0745	0·000745
		1·00	4	0·451	0·00451
			5	1·74	0·0174

2.5. *The electrostriction force*

Electrostriction is defined as the elastic deformation of a dielectric under the force exerted by an electric field. A dielectric can be imagined to consist of dielectric particles (molecules) uniformly distributed throughout the vacuum space. Thus the internal field acting on such particles is higher than the apparent or measured field strength E by an amount resulting from the polarization field of the polarized dielectric particles. This internal field (Bottcher 1952) is

$$E_i = \frac{\epsilon + 2\epsilon_0}{3\epsilon_0}E \quad (26)$$

where ϵ and ϵ_0 are the permittivities of the dielectric and vacuum respectively. From Eqn (1) the decrease in stored energy per unit volume due to the presence of such particles in the vacuum space is

$$\frac{dU_v}{dv} = \frac{\epsilon_0 E^2}{2}\frac{(\epsilon/\epsilon_0 + 2)(\epsilon/\epsilon_0 - 1)}{3}. \quad (27)$$

Hence the electrostriction force per unit volume is

$$F_v = \frac{\epsilon_0(\epsilon/\epsilon_0 + 2)(\epsilon/\epsilon_0 - 1)}{6}\nabla E^2 \quad (28)$$

and the mechanical stress tending to contract the dielectric is

$$P_v = \frac{\epsilon_0(\epsilon/\epsilon_0 + 2)(\epsilon/\epsilon_0 - 1)}{6}E^2. \quad (29)$$

It is obvious that in a uniform field there is no translational force, but in a non-uniform field this force will tend to push the particles (molecules) from the weak to the intense field. In gases this effect may be negligible because their relative permittivities are approximately unity, but in solid or liquid dielectrics this effect may be important under a highly divergent field. For example, at a field of 1 MV/cm the mechanical pressure contracting the water whose relative permittivity is about 80 is about 13 800 lb/in^2. This is an extremely high pressure and may be of interest to practical applications.

2.6. *The torque orientating a solid body*

If a body is not symmetrical about its centre, it will experience a torque tending to orientate itself in such a manner that the stored energy in the field will become minimum. For simplicity a prolate spheroidal shape is chosen to be the approximation to the general shape of the solid body, and its principal axes, which follow the relationship $a > b = c$, form a rectangular coordinate frame, X directing the major axis, $2a$, Y and Z the minor axes, $2b$ and $2c$. If one of the minor axes is originally in the direction of the field, the decrease in stored energy when its major axis orientates and causes an angular displacement θ, from Eqns (1) and (12), is

$$U_T = \frac{2\pi ab^2\epsilon_1(\epsilon_2 - \epsilon_1)^2(G_y - G_x)\sin^2\theta}{3\{\epsilon_1 - (\epsilon_1 - \epsilon_2)G_x\}\{\epsilon_1 - (\epsilon_1 - \epsilon_2)G_y\}}E_1^2. \quad (30)$$

Hence the torque exerted on the body is

$$T = \frac{dU_T}{d\theta} = \frac{2\pi ab^2\epsilon_1(\epsilon_2 - \epsilon_1)^2(G_y - G_x)\sin 2\theta}{3\{\epsilon_1 - (\epsilon_1 - \epsilon_2)G_x\}\{\epsilon_1 - (\epsilon_1 - \epsilon_2)G_y\}}E_1^2 \quad (31)$$

where

$$G_x = \frac{ab^2}{2}\int_0^\infty \frac{ds}{(s + a^2)^{3/2}(s + b^2)}$$
$$= -\tfrac{1}{2}\frac{(b/a)^2}{\{1 - (b/a)^2\}^{3/2}}\left[2\{1 - (b/a)^2\}^{1/2} + \log\frac{1 - \{1 - (b/a)^2\}^{1/2}}{1 + \{1 - (b/a)^2\}^{1/2}}\right] \quad (32)$$

$$G_y = \frac{ab^2}{2}\int_0^\infty \frac{ds}{(s + a^2)^{1/2}(s + b^2)^2}$$
$$= \tfrac{1}{4}\frac{(b/a)^2}{\{1 - (b/a)^2\}^{3/2}}\left[\frac{2\{1 - (b/a)^2\}^{1/2}}{(b/a)^2} + \log\frac{1 - \{1 - (b/a)^2\}^{1/2}}{1 + \{1 - (b/a)^2\}^{1/2}}\right]. \quad (33)$$

E_1 is the field strength in the region of the dielectric medium where the body is subsequently present and ϵ_1 and ϵ_2 are the permittivities of the medium and the body respectively.

The torque tends to orientate the major axis of the body to the direction of the field and is independent of the relative magnitudes of ϵ_1 and ϵ_2. The magnitude of the torque depends on the size and the position of the body. It is zero when the major axis of the body is perpendicular to the field, and is maximum when it is at the angle of 45° to the field. The torque exerted on the solid particles may accentuate the effect of impurities on the dielectric breakdown, but on the other hand, it may be used as a means to determine the permittivity of a dielectric sample of known shape.

3. Conclusions

The formulae for the calculation of the mechanical forces which result only from polarization have been derived. The magnitudes of these forces may be appreciable when the impressed electric fields are large. The directions of these forces are independent of the directions of the fields, that is, either unidirectional or alternating fields will produce the same effects.

Acknowledgments

The author wishes to thank the Directors of the Brush Electrical Engineering Company Limited for permission to publish this paper.

References

BOTTCHER, C. J. F., 1952, *Theory of Electric Polarization* (New York: Elsevier Co.).

COOPER, R., and WALLACE, A. A., 1953, *Proc. Phys. Soc.* B, **66**, 1113.

KAO, K. C., and HIGHAM, J. B., 1961, *J. Electrochem. Soc.*, **108**, 522.

POHL, H. A., 1958, *J. Appl. Phys.*, **29**, 1182.

STARK, K. H., and GARTON, C. G., 1955, *Nature*, **176**, 1225.

STRATTON, J. A., 1941, *Electromagnetic Theory* (New York: McGraw-Hill).

New Books

Introduction to the dynamics of automatic regulating of electrical machines. By M. V. MEEROV. (London: Butterworths, 1961.) Pp. xxii + 411. Price 100*s*.

If a technical book published in the Soviet Union has special merit in a particular field or covers a subject not well documented in the English language, the National Lending Library for Science and Technology, Department of Scientific and Industrial Research may arrange for its translation and publication in the U.K. This book, first published in Russian in 1956, is such a work and J. S. Shapiro is the translator. Let us say at once that an excellent job has been made by both translator and publisher: the text reads without any trace of having been translated and the type-setting throughout is good, in text, mathematics and diagrams.

Many automatic systems depend, in great part, on automatically regulated electrical drives, and generating systems depend for effective operation on the close control of output voltage and frequency. These types of closed-loop regulating systems are discussed in detail by the author and the whole matter is laid out logically and clearly.

There are three basic problems in any automatic regulating system and these are concerned with stability, quality and synthesis. Whilst it is well to consider each separately, yet there is a good deal of interdependence, and this is brought out very clearly as the text proceeds.

The early chapters deal with the characteristic equations and the time constants of the basic components of regulating systems: full consideration is given to such items as d.c. and a.c. generators, d.c. motors, tachogenerators, rotary electrical amplifiers (amplidyne, etc.), magnetic and electronic amplifiers, synchros, etc. Problems of stability and stabilizing devices, and then of quality are next dealt with, after which the synthesis of regulating systems is discussed. In all of these subjects examples are given and a separate chapter deals with three specific examples, two for governing the speed of a d.c. motor under specified conditions, and a third on the voltage regulation of two a.c. generators connected in parallel over a long transmission line and supplying a load. Separate chapters are devoted to intermittent regulating systems and to non-linear systems.

As might perhaps be expected all references in the bibliography are to Russian papers but the text is so complete in itself that there should be little need to want access to them. A list of ten books in English on control systems and principles is appended.

The book was well worth translating into English and we consider it should be of great value to practising engineers and others concerned with the technology of automatic control engineering and electrical drive systems.

A. J. MADDOCK

Compensating radiometer

by A. R. KING, Ph.D., Division of Physical Chemistry, Commonwealth Scientific and Industrial Research Organization, Chemical Research Laboratories, Melbourne, Australia

MS. received 20*th June* 1961, *in revised form* 8*th August* 1961

Abstract

A radiometer for the measurement of near infra-red radiation in fire research is described. The novel features of the instrument are robustness, ability to withstand immersion in flames and the absence of windows.

Introduction

RESEARCH on forest fire has been hampered by the lack of a suitable radiometer for use in the field near rapidly moving flame fronts and for some laboratory applications. This paper describes an instrument that has been devised to fill this gap and that has been proven over two fire seasons in Australia.

General requirements

To collect radiation intensity–time relationships the instrument is sited in the path of the fire and connected by an asbestos-insulated cable to a potentiometer some distance away. Readings are taken until the flames reach the instrument which is then either withdrawn or left to withstand the flames. This severe use calls for special design and the criteria for a suitable instrument are as follows:

(i) it should be sufficiently robust to stand up to rough usage and vibration during transport;
(ii) it should be small, uncomplicated and easily set up;
(iii) it should be cheap, since fairly frequent loss or destruction is unavoidable;
(iv) it should be undamaged by exposure to radiation of 0·5 cal cm^{-2} s^{-1} for 15 minutes and by immersion in the usual diffusion flames for 5 minutes;
(v) its acceptance angle should be 180° and the receiving element should follow the cosine rule of absorptivity;
(vi) its range should be 0–1 cal cm^{-2} s^{-1};
(vii) its response should be non-selective over the wavelength range of 0·4–12 μ, stable and preferably linear;
(viii) its rate of response should exceed 90% of change in 10 seconds;
(ix) its sensitivity should not be greatly affected by changes in air temperature (20–200° C), wind speed (0–7 miles/h), wind direction relative to the instrument and the temperature of the instrument itself (20–100° C);
(x) its accuracy should be ±2% under laboratory conditions. A lower accuracy is satisfactory for field work where conditions are very variable, the fire intensity is constantly fluctuating and the related data such as flame height and rate of spread cannot be measured very accurately.

Description of instrument

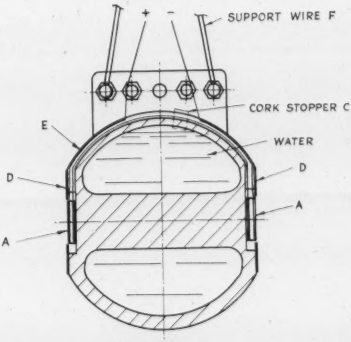
Most instruments already developed (Gergen 1956, Gier *et al.* 1951, Lowry 1957, Marshall and MacKenzie 1950, McGuire and Wraight 1960, Suomi 1954) fail because they are not sufficiently robust and cannot be fireproofed conveniently. Also, most need to be shielded against the outside air with windows which must be kept scrupulously clean, (Gergen 1956, Lowry 1957, Marshall and MacKenzie 1950, McGuire and Wraight 1960, Suomi 1954) that are usually fragile and costly (Gergen 1956, Marshall and MacKenzie 1950, Suomi 1954) or of low melting point (Lowry 1957). Calorimeters (Butler 1953) could be used but they would not discriminate between convected and radiated heat unless suffering from many other disadvantages.

The operating principle of the radiometer devised is novel, it being to measure the heat absorbed by two identical planes, one of which is subjected to radiated and convected heat and the other to convected heat alone. The measurements are then automatically subtracted to give an output proportional to the radiation intensity alone.

The planes which receive the heat are faces of two matched heatflow meter disks described by Hatfield and Wilkins (1950). These are disks (1 mm thick × 1 cm diameter) of a silver–tellurium (1 : 2) alloy with circles of copper gauze attached thermally and electrically to each planar face. The disks are used in such a way that the heat flux to be measured flows in through one gauze circle and out through the other into a heat sink. A temperature gradient is set up proportional to the rate of flow of heat and this in turn generates a proportional e.m.f. difference between the two junctions, gauze to alloy, which constitute thermocouples. The e.m.f. difference is tapped by wires attached to the copper gauzes.

Precautions that should be observed are that the temperature of the disks should not exceed 100° C for continuous use and 150° C intermittently, that the faces are uniformly exposed to, or in contact with, the heat source and sink and that heat does not flow in or out of the edge of the disks.

The details of the radiometer are shown in the Figure. The disks A are glued with rubber cement to the bottom of two cylindrical holes (0·05 in. deep × ½ in. diameter) machined into opposite sides of a brass sphere. This sphere is made in two halves, which are soldered together, and has an annular cavity which is filled with water through a small



hole C stoppered with a cork. The sphere acts as a heat sink, minimizes rapid temperature fluctuation and ensures that the inner faces of the two disks are essentially at the same temperature. If the instrument is heated strongly, the water boils, ejecting the cork and, for a considerable time at least, prevents the temperature from exceeding 100° C.

There is an air space around the edges of the disks but the shields D, which lie in the plane of the outer gauze of the disks and almost touch them, prevent radiation from falling on the edges.

The inner gauzes of the disks are electrically connected by a wire running along a groove E cut around the sphere. Wires connecting the outer gauzes to the potentiometer also run along the groove past the instrument support F.

The disk intended to absorb the convected and radiated heat, and which consequently is pointed towards the flames, is painted with 'Blackboard Black' paint produced by Taubmans Pty. Ltd., Australia. The absorptivity of this paint even after misuse and rubbing is 90–95% as determined by comparison with a thick (0·2–0·5 mg/cm^2) coating of fine carbon black deposited from a small camphor flame. Kozyrev and Vershinim (1959) have shown that such coatings have an absorptivity of 95–98%. In field work, the coating is renewed when necessary but care must be taken to ensure that the paint is dry before use. The radiometer receiving face is also coated with a similar deposit of carbon black for most accurate measurements.

The disk intended to absorb only convected heat is covered with a circle of bright aluminium foil stuck on with high melting-point grease such as silicone stop-cock grease. The absorptivity of the foil is less than 5% and when eventually dirtied it can be replaced easily. Normally, the face is shielded from the radiation of the fire by the physical obstruction of the sphere but the foil is used because, in some cases, there may be some flame or hot material within its angle of view.

Calibration and recovery

The instrument was calibrated using a constant-flow water calorimeter as used at the Fire Research Station of the Department of Scientific and Industrial Research, England, and the Fire Section of the National Research Council of Canada. This calorimeter was coated using a camphor flame after painting with 'Blackboard Black' paint as a double check to ensure complete absorption. The temperature rise observed when subjected to radiation was equalled by heating with known electrical power.

The calibration of the radiometers is almost linear and the outputs approximate 18 mv cal^{-1} cm^2 s. The reproducibility of the calibration was found to be better than ±2% and even so errors could have been largely introduced by the auxiliary equipment.

The speed of response is high for a field instrument, being 10 seconds for 95% full scale when using a stepwise increase or decrease of radiation of 0·1 cal cm^{-2} s^{-1}.

Within the experimental error the carbon black and the paint used on the radiometers were found to follow the cosine law as shown in the Table. The best source available for these checks was a conventional electrical infra-red lamp which inevitably introduced errors since rays from it are not parallel.

In still air, the error introduced by heating the instrument to 100° C (i.e., 75 deg C above the ambient temperature) was +6% at a radiation intensity of 0·1 cal cm^{-2} s^{-1}. When the air temperature exceeded that of the instrument by 75 deg C the error was +4% at a radiation intensity of 0·1 cal cm^{-2}s^{-1}.

Cosine law

Angle of incidence, degrees	0	6	21	36	51	66
Value corrected for angle (cal cm^{-2} s^{-1})						
Carbon black	0·050	0·050	0·050	0·049	0·052	0·054
Paint	0·048	0·048	0·048	0·045	0·051	0·053

When winds of 4 miles/h were directed on to the blackened face when it was 10 deg C colder than the air and the radiation intensity was 0·2 cal/cm^2 s the error was −10%. Under the same conditions, the error was +6% if the wind direction was reversed and +1% if it blew parallel to the faces.

Errors may soar to 50% under the worst conditions as, for example, with a wind of 7 miles/h blowing on to the blackened face when the temperatures of air and radiometer differ by 75 deg C. This wind speed is rare in forests and the maximum temperature difference before the instrument is immersed in the flame has been found to be 15 deg C. (This figure may seem low to those acquainted with large fires but is based on many measurements made with specially devised shielded thermocouples and aspirated thermometer adjacent to large flame fronts.)

When flames lick the instrument the readings become very variable. Unfortunately there did not seem to be any way to check the accuracy of the instrument under these conditions.

To summarize the effect of wind and temperature, the errors to be expected in usual field measurements will be less than ±10% and will decrease with decreasing wind, with wind direction approaching the planar direction of the receiving faces and with decreasing difference in temperature between the instrument and the surrounding air. In normal draughty laboratories, the errors will be less than ±2%.

The electrical resistance of the radiometers constructed is less than 1 ohm.

Other uses

In addition to normal use the radiometers have been used to measure the reflectivity of a variety of pigments, paints, barrier creams, fabrics and metallic foils. The technique is to apply the material under test to both receiving elements of a radiometer; pigment is dusted on to a grease, paints and creams are brushed on and fabrics and foils are stuck on with grease. Application of the material to both faces ensures reasonably equal convective transfer at each. The radiometer is then used in the usual fashion and the ratio of the difference between the actual radiation intensity and the absorption measured to the actual radiation intensity equals the reflectivity of the material.

References

BUTLER, C. P., 1953, *U.S. Nav. Radiol. Def. Lab. Rep.*, 397 (Physics).

GERGEN, J. L., 1956, *Rev. Sci. Instrum.* **27,** 453.

GIER, J. F., *et al.*, 1951, *Trans Amer. Inst. Elect. Engrs*, **70,** 339.

HATFIELD, H. S., and WILKINS, F. J., 1950, *J. Sci. Instrum.*, **27,** 1.

KOZYREV, B. P., and VERSHINIM, O. E., 1959, *Optika J. Spekroskopii*, **6** (4), 345.

LOWRY, W. P., 1957, *Ecology*, **38,** 152.

MARSHALL, P. R., and MACKENZIE, D. K., 1950, *J. Sci. Instrum.*, **27,** 33.

MCGUIRE, J. H., and WRAIGHT, H., 1960, *J. Sci. Instrum.*, **37,** 128.

SUOMI, V. E., 1954, *J. Met.*, **11,** 276.

Influence of the electric field on the silver and iron whiskers growth

by T. HOFFMANN,* M Sc., J. MAZUR,† D.Phil., F.Inst.P., J. NIKLIBORC,* D.Phil., and J. RAFAŁOWICZ,† M.Sc.

* Physics Department, The University, Wrocław, Poland; † Physics Department, The University, and Low Temperature Laboratory, Institute of Physics, Polish Academy of Sciences, Wrocław, Poland

MS. received 10*th August* 1961

Abstract

The influence of an electric field on the silver and iron whiskers growth obtained by Brenner's method has been studied. There is a marked directional tendency during the growth process; the findings are compared with results of investigation on copper.

Introduction

THE present work is a continuation of the studies already published (Hoffmann *et al.* 1961) concerning the influence of the electric field on the copper whiskers growth; an ample list of references was included.

In this paper we describe the results of the investigation of the influence of the electric field on the silver and iron whiskers growth. The experiments have been carried out by using the apparatus already described in detail in our previous paper.

The silver whiskers have been obtained by hydrogen reduction of silver chloride and iron whiskers by hydrogen reduction of ferrous chloride. Silver electrodes have been used for production of the silver whiskers and iron electrodes for iron whiskers. The distance between electrodes was 20 mm. Temperature of the reduction was 800° C for silver and 750° C for iron. The applied electric field did not exceed 1000 V.

Results

The results of the investigation may be summarized as follows.

Silver whiskers

(i) The electric field has a marked influence both on the formation and on the direction of the whiskers.

(ii) The whiskers produced in the presence of an electric field are longer and more numerous than those obtained without field.

(iii) The majority of the whiskers grown on the lower electrode, when the electric field has been applied, are ramified, sometimes manifoldly (Fig. 1).

(iv) The majority of whiskers grown on the lower electrode, when the electric field is applied follow the direction of the field lines. There is no preferred direction of the whiskers grown in the absence of the electric field.

(v) On the upper electrode whiskers grow, when the electric field is applied, irrespective of the direction of the field. During the absence of the electric field no whiskers have been observed. The whiskers on the upper electrode follow the field lines in more striking manner than those on the lower boat (Fig. 2).

Fig. 1. Silver whiskers on the lower electrode. ×100.

In a few cases the ramification of the whiskers on the upper electrode have been confirmed (Fig. 3).

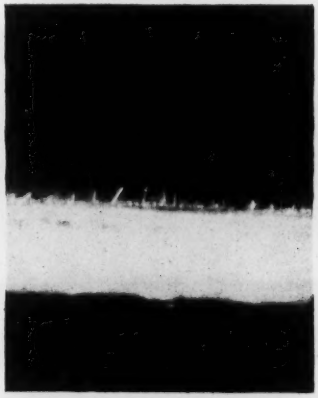


Fig. 2. Silver whiskers on the upper electrode. ×100.

Fig. 3. Some ramified whiskers on the upper electrode. ×100.

Iron whiskers

The influence of the electric field on the growth of the iron whiskers is exactly the same as on the silver whiskers; the only one difference consists in the fact that the iron whiskers grown on the upper electrode are not so numerous, nor so long, as silver whiskers (Fig. 4).

Conclusions

We conclude that, by applying Brenner's method, we can most easily grow copper whiskers on the upper electrode but that silver whiskers are easier to grow than iron.

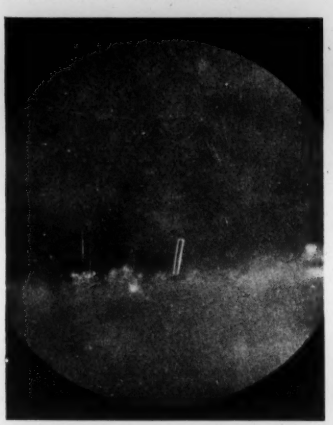


Fig. 4. The iron whiskers on the lower electrode. Notice a very interesting specimen of the whisker in a form of a rectangle. ×100.

As far as the electric field is concerned our present findings confirm entirely the results of the study on copper.

The investigation on the mechanism of the production of the whiskers when the electric field is applied, is in progress.

Reference

HOFFMANN, T., MAZUR, J., NIKLIBORC, J., and RAFAŁOWICZ, J., 1961, *Brit. J. Appl. Phys.*, **38**, 342.

Method of measuring the growth of a control specimen of graphite after irradiation

by P. A. E. CROSSE and W. L. SNOWSILL, B.Sc., A.Inst.P., Central Electricity Research Laboratories, Kingston Road, Leatherhead, Surrey

MS. received 3rd July 1961

Abstract

A method is described for measuring the length of control specimens of graphite accurately without placing any reliance on the condition of the surfaces of the graphite.

A specially shaped groove is turned near each end of the 10 cm long × 1·25 cm diameter specimen, and is filled with flame-sprayed Magnox. A technique is described for producing a high-quality surface on the Magnox on which a diamond indentation under microscopic examination appears as a pair of cross wires.

Using the 'cross wires' as datum marks, and a high-quality travelling microscope, measurements to an accuracy of 1 in 25000 are readily attainable. Methods are discussed for increasing this to a possible 1 : 100000.

1. Introduction

It is a practice in nuclear reactors to include somewhere, usually in one or more of the fuel element channels, a number of specimens of graphite, called control specimens, which can be withdrawn periodically and examined for changes in physical properties, thus monitoring the condition of the whole mass of graphite.

Graphite, when irradiated, changes its dimensions, and accurate measurement of the growth of the control specimens is essential to provide information about the dimensional changes of the different parts of the moderator.

Growth is usually measured by what is potentially an extremely accurate method. A one inch growth cube, machined and ground to optically fine limits is measured by comparison with slip gauges on a Sigma comparator. Measurement to an accuracy of 1 : 250000 is possible by this method, but it has the following disadvantages:

(i) Since graphite is transferred from a surface in one part of the reactor to another at a different temperature, using carbon dioxide as a carrier, a build-up or a reduction of graphite can occur on the finely worked surfaces, giving a false indication of dimensional changes.

(ii) The preparation of the specimens is very expensive.

The following method of measuring the change in length does not involve the use of the end faces, and an accuracy of 1 : 25000 is obtainable.

2. Description of the method

The simplest method of measuring the change in length of a specimen, is to measure the change in the difference between two datum points near the extremities of the specimen using a travelling microscope. The specimens are 10 cm long and 1·25 cm diameter, so that for the orders of accuracy required, measurements to 1 micron are necessary.

Since the surface of graphite is likely to deteriorate during prolonged service in the reactor for the reasons mentioned earlier, it is not practicable to cut datum marks in the graphite itself. A better method is to prepare metal surfaces on the graphite and to place datum marks in them. The metal must be compatible with Magnox and must not absorb neutrons. This is the principle of the new method and the metal chosen is Magnox itself.

A groove is cut into the specimen at each end, as in Fig. 1(*a*). Pure Magnox is flame-sprayed into the groove,

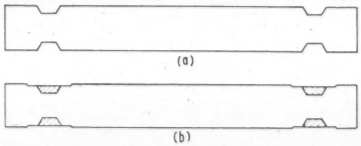


Fig. 1. (*a*) Specimen before spraying. (*b*) Specimen after spraying and machining.

until the ring so formed is larger in diameter than the specimen. This ring is finely turned using a diamond cutting tool, and the turned surface then rolled using a polished stainless steel roller. The surface produced by this method does not exhibit the usual furrows produced by turning on a lathe, even under a microscope at a magnification of ×100. As is seen in Fig. 1(*b*) the diameter of the ring after turning is less than that of the main body of the graphite to reduce the likelihood of surface damage in handling.

Each of the above processes needs to be carried out in clean conditions to prevent contamination of the graphite. While this presents no difficulty as far as the diamond indenting is concerned, care is necessary in the flame spraying operation. Exhauster fans which are required to remove the excess magnox dust, create draughts which tend to introduce contamination. It is also possible for impurities to enter via the gases supplying the flame gun. This latter is minimized by using well filtered compressed air and fresh cylinders of oxygen and acetylene. Most of the impurities introduced by the flame spraying process are removed in the final machining when the diameter of the specimen is reduced by a further 1 mm.

Four pyramidal indentations are then made in each ring, equally spaced around it, using the diamond conveniently available in a Vickers hardness tester. The diamond marks on each ring are aligned to enable them to be used as datum

marks for length measurement, four being required to provide a check on warping of the specimen. The Vickers hardness tester is loaded to 15 kg so that the indentations are deep enough to reduce still further the possibilities of damage during handling. The pyramid shaped indentations, under a microscope with suitable illumination, have the appearance of cross wires, but in a very much more precise form than is possible using scribing techniques.

The distance between pairs of 'cross wires' is measured using an accurate travelling microscope, with a magnification of $\times 40$. This has a micrometer adjustment calibrated in 10-micron divisions, with a vernier scale enabling measurement to 1 micron. A limitation of the standard microscope is the difficulty of aligning the normal cross-wire graticule on the extremely fine diamond impression. This has been improved by replacing the standard graticule with one consisting of a pair of parallel wires, nominally 100 microns apart. A reading is taken by aligning the wires so that they evenly straddle one of the diagonals of the diamond impression (see Fig. 2). By this method independent operators can repeat readings to an estimated 0·5 micron on the vernier.

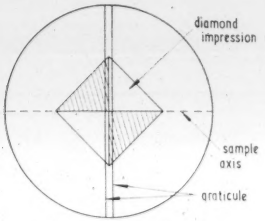


Fig. 2. Appearance of datum marks under microscope.

It can be seen that placing a ring of one coefficient of expansion in a groove of a material with a different coefficient of expansion may lead, when the temperatures rises, either to stress in the materials or to loosening of the ring, unless the groove shape is so designed to prevent this. It is important that the ring should not become very loose because this would allow graphite to be deposited on, or removed from the groove by the process discussed earlier. It would also be possible for graphite dust to collect in the spaces.

The correct groove shape has been calculated theoretically with reference to Fig. 3 as follows. The following assumptions have been made:

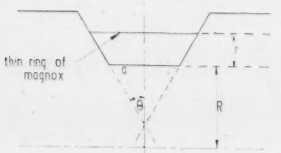


Fig. 3. Development of groove shape.

Magnox, in the sprayed form, is isotropic with a coefficient of expansion α. Graphite is anisotropic, and has coefficients of expansion β_1 longitudinally, β_2 transversely in the plane of the paper, β_3 transversely normal to the plane of the paper. Let θ be the angle between the groove and the normal to the axis, $2a$ the length of the bottom of the groove, and R the radius of the bottom of the groove.

Consider a ring of Magnox, radius $R + r$. The necessary conditions for the ring to remain in contact with the groove walls as the temperature changes are

$$\tan\theta_1 = \frac{(a + r\tan\theta_1)(1 + \alpha t) - a(1 + \beta_1 t)}{(R + r)(1 + \alpha t) - R(1 + \beta_2 t)}$$

for a segment of ring in the plane of the paper

i.e. $$\tan\theta_1 = \frac{a(\alpha - \beta_1)}{R(\alpha - \beta_2)}$$

and $$\tan\theta_2 = \frac{a(\alpha - \beta_1)}{R(\alpha - \beta_3)}$$

for a segment of ring in the plane normal to the paper.

Now, if $\beta_1 = \beta_2 = \beta_3$, $\tan\theta_1 = \tan\theta_2 = a/R$. A typical value for α is 28×10^{-6} per deg C between 20° C and 400° C. Typical values for β vary between $2{\cdot}4 \times 10^{-6}$ per deg C parallel to the extrusion axis and $4{\cdot}2 \times 10^{-6}$ per deg C perpendicular to the extrusion axis. In practice, the difference between β_1, β_2 and β_3 is sufficiently small compared with the difference between α and β to say $\tan\theta = a/R$ and the edges of the ring remain substantially in contact at all temperatures. Obviously, unless $\beta_2 = \beta_3$ machining the perfect groove would be extremely difficult. For specimens cut with the axis parallel to the extrusion axis, β_2 and β_3 are equal.

3. Thermal cycling experiments

Three specimens produced by the above method were heated to 420° C and cooled to room temperature in an atmosphere of carbon dioxide, twenty times. Each time the specimens were kept at the upper temperature for a period of approximately 15 minutes, and the heating and cooling cycles took approximately 45 minutes, and 30 minutes respectively.

Measurements were taken of the length of the specimens at the end of many of the cycles, and the maximum variation on any of the specimens was within ± 2 microns, i.e. over a length of 10 cm, 1 : 25 000. Possible reasons for the spread of ± 2 microns rather than ± 1, as indicated above, are discussed later. Apart from a slight dulling of the Magnox surface, the specimens were unaltered in appearance, and the accuracy of setting was unchanged.

4. Limitations of the method

Whereas it is possible to repeat readings on a given diamond impression to within 0·5 micron, the thermal cycling experiments showed that length measurements are repeatable only to an accuracy of ± 2 microns.

When the Magnox rings are sprayed on to the graphite, the temperature of the specimen rises to approximately 60° C. At this temperature the ring is a perfect fit in the groove, both on the sides and the base. Should the temperature rise from this point, the ring would remain in contact with the sides of the groove, as indicated in the previous theoretical analysis, and move away from the base. On cooling below this temperature however, even down to room temperature, a strain is bound to occur. The ring tightens on to the base of the groove and relaxes from the sides. The differential expansion coefficient between Magnox and graphite is of the order of 22×10^{-6} per deg C which, with an angle $\theta = 25°$, and $R = 3$ mm gives a differential shrinkage longitudinally of approximately 3 microns when cooling down to room

temperature. It is therefore feasible that above room temperature each ring will be capable of a movement longitudinally of 3 microns giving a possible length error of ± 3 microns. In fact this gap is probably reduced due to the general straining of the ring at the base of the groove, so that the errors produced during thermal cycling could be explained by the movement of the ring relative to the specimen, when hot. The amount of possible movement could be decreased by reducing θ and a, but this in fact impairs the structure of the ring as many voids are introduced into the Magnox during spraying if θ is less than about 25°. However, it is considered that the gap involved is sufficiently small to be of no consequence as far as graphite transfer is concerned.

A further series of tests had been proposed to offset the strain produced on cooling following spraying. Suggested methods are:

(i) the inclusion of a 'V' cut or a 'V' projection in the base of the groove to locate the ring so that at room temperatures it will always return to the same spot,

(ii) the inclusion of a material in the base of the groove before spraying, which is capable of giving sufficiently on initial cooling to prevent strain.

This could increase the accuracy to the maximum possible with this method, of $\pm\frac{1}{2}$ micron in 10 cm, i.e. in the region of 1 in 100 000.

5. Conclusions

It is possible to measure the length of a specimen of graphite to an accuracy of 1 in 25 000 without placing any reliance on the condition of the end faces or in fact any exposed surface of the graphite. The method would seem to be inherently capable of an accuracy of 1 in 50 000 with a slightly modified groove, and even 1 in 100 000 given skilled operators.

Acknowledgments

This article is published by permission of the Director, Central Electricity Research Laboratories.

New Books

X-ray metallography. By A. Taylor. (London: John Wiley, 1961.) Pp. vii + 993. Price 216*s*.

In a word, this is a more comprehensive version of the book by the same author 'An introduction to x-ray metallography' (Pp. xi + 400) published by Chapman and Hall in 1945. The layout bears a strong resemblance to that of the earlier book but there has been a very extensive revision of the text both to bring it up to date and to include some new topics not dealt with previously.

The main theme of the book is the application of x-ray techniques to the study of metals and alloys. It is directed both to the student reading for a degree and to research workers. Although it is a practical book with a wealth of detail on the application of x-ray techniques, there is a very full and lucid treatment of the theoretical background to the subject. In fact, a feature which will please the serious student is the advanced level to which the theory is taken. Thus, in the section on Crystal Symmetry one finds a logical development of the subject up to and including the concept of space groups. Many metallurgical texts contain little more than a list of definitions here.

About one quarter of the whole is devoted to this introduction to crystallography, the generation and properties of x-rays and the principles of x-ray structure analysis. Then follows the real meat of the book which we may conveniently divide into:

1. The study of thermal equilibrium diagrams: alloy chemistry, and
2. The size, perfection and orientation textures of grains in polycrystalline aggregates.

These topics span several chapters. Two of them where special attention has been given to recent developments are: 'Precipitation hardening, complex alloys, and steels' and 'Internal Stresses in Metals and Alloys'.

Fluorescence analysis is described with a good deal of practical detail in the new chapter on 'Chemical analysis by x-ray procedures'. The electron probe microanalyser receives rather brief attention.

To sum up: this enormous work which makes reference to papers published into the late 1950's and which includes 88 useful tables in the Appendix, will prove a great asset to the research man. The undergraduate who is fortunate to lay hold on it will find his searching well rewarded too. Despite the price it will prove a worthy rival to other texts in this field.

R. L. BELL

Notes and comments

Notes for the Preparation of Papers for the *British Journal of Applied Physics*

The following instructions refer to the main points which have to be considered in the preparation of a paper for publication, and authors offering papers to the *British Journal of Applied Physics* for publication are asked to conform to these recommendations.

Manuscripts

Manuscripts should be typed in *double* spacing on paper not wider than 8 in. and not longer than 13 in. Only one side of the paper should be used, and a margin of 1–1½ in. should be left. As alterations in the text cannot be allowed once the paper is set up in type, authors should aim at absolute clarity of meaning and of typing, and should check the typescript carefully before submission. All manuscripts should be submitted in duplicate, and in addition to the fair copies, a set of small copies or prints of the diagrams (not larger than foolscap) must be attached to each MS. (This enables MSS. to be sent to two referees at the same time and so assists rapid publication.)

Abstract. An abstract is printed at the beginning of the paper immediately after the title, name(s) of author(s), and place of employment of author(s). Two extra copies of the 'title and abstract' page are required for the records.

Mathematics. It is not necessary to give detailed derivations of mathematical expressions and formulae in a published paper when the work is straightforward; it is quite sufficient to indicate the method of treatment and the final results.

References. In the text bibliographical references are made by giving the name of the author and the year of publication in brackets, e.g. (Jones 1942), and details are given in the last section, 'References', where the references are arranged in alphabetical order of authors' names and in date order for each author.

Drawings

Drawings should be in Indian ink on tracing cloth, tracing paper or white card, with lettering in soft or blue pencil; lettering of a size suitable for reduction will be inserted by our draughtsman. The drawings should in general be sufficiently large to allow of reduction in printing, and the lines should therefore be bold; the frame lines of graphs should be slightly finer than those of the plotted curves.

Full details are available in the Institute and Society's 'Notes for Authors', obtainable from the Editor and Deputy Secretary, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London S.W.7.

Papers for publication in the *British Journal of Applied Physics* should be sent to the Editor and Deputy Secretary at the same address.

Annual Exhibition of Scientific Instruments and Apparatus

The Annual Exhibition of The Institute of Physics and The Physical Society will be held on 15th–19th January, 1962, at the Royal Horticultural Society's Halls, Westminster, London S.W.1. The latest developments in British scientific equipment will be shown.

Further information may be obtained from The Institute of Physics and The Physical Society, 47 Belgrave Square, London S.W.1.

Weizmann Memorial Fellowships 1962–63

The Weizmann Memorial Foundation will shortly elect four Fellows to spend a year in research in the natural sciences at the Weizmann Institute of Science, Rehovoth, Israel, beginning in the autumn of 1962. These fellowships are intended for scientists with several years of post-doctoral research experience.

Further information may be obtained from the Academic Secretary, The Weizmann Institute of Science, Rehovoth, Israel, and applications must be received not later than 1st December, 1961.

Journal of Scientific Instruments

Contents of the November issue

CONFERENCE REPORTS

Conference on optical instruments and techniques—London, 11th–14th July 1961. By W. J. Bates.

Colloquium on optical materials—Paris, 1961. By K. M. Greenland.

PAPERS

Instrument for the measurement of elastic hysteresis in torsion. By R. Diamond and J. E. Whittle.

Analogue computer for the summation of Fourier series with digital recording of the results. By T. Lundström and V. Klimecki.

Radial shearing interferometer. By P. Hariharan and D. Sen.

Measurement of the figure of merit of a thermoelectric material. By A. E. Bowley, L. E. J. Cowles, G. J. Williams and H. J. Goldsmid.

Miniature cryostat for measurements on ferromagnetics between 1·6 and 300° K. By W. R. Scott and J. Crangle.

High-pressure hydrogen gas target. By J. Kirk.

Height of waves in hydraulic models by the method of stereoscopy. By J. W. Railly and B. F. Scott.

Recording refractometer for measuring refractive indices of optical glasses over a wide range of wavelengths. By M. J. C. Flude, K. J. Habell and A. Jackson.

Humidity control device for the Buerger precession camera. By J. R. Einstein.

LABORATORY AND WORKSHOP NOTES

Heat cycling apparatus for thermostat testing. By T. D. Bansal and S. K. Chhabra.

Improved reflection diffraction holder for electron microscope. By J. C. Henderson.

NOTES AND NEWS

New instruments, materials and tools New books Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique).

All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.

REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611–2.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12*s.* 6*d.* each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.

CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

PRINTED IN GREAT BRITAIN BY UNWIN BROTHERS LIMITED, WOKING AND LONDON

Rutherford Jubilee International Conference—Manchester, September 1961

Abstract

A conference sponsored by The International Union of Pure and Applied Physics, The Royal Society, The Institute of Physics and The Physical Society, and The Victoria University of Manchester was held at Manchester between the 4th and 8th of September 1961, in recognition of the jubilee of the discovery of the atomic nucleus by Rutherford. This arose directly from the experiments of Geiger and Marsden in Rutherford's laboratory at Manchester. Under the presidency of Sir Ernest Marsden a wide range of topics in experimental and theoretical nuclear physics was covered by a number of invited speakers. A special commemorative session was included in the proceedings.

Physics, despite its forbidding appearance to many of its students, is a creative human activity and the history of physics is largely a story of the work of great men. Within about a century, four of the greatest in science, namely Dalton, Joule, Thomson and Rutherford, lived in the city of Manchester. Their work has given us our modern picture of the atom and ultimately our control of its innermost energies. The atomic nucleus, in which the most powerful forces of the atom reside, was conceived by Rutherford to explain the scattering of α particles by heavy elements, and was proposed in a classic paper of compelling simplicity in the Philosophical Magazine for May 1911. Fifty years later the University of Manchester, in which Rutherford held the Langworthy Chair of Physics, has marked the jubilee of the most important concept of modern physics by holding a large International Conference to which, despite its size, only a small fraction of the nuclear physicists of the world could be invited. Among them, most happily, was Niels Bohr, whose brilliant interpretation of the hydrogen spectrum in terms of the nuclear atom (1913) originated from his stay in Manchester. The Bohr theory of the atom, even more than the evidence of radioactivity, forced the nucleus into physics.

The conference was fortunate in being able to appoint as President Sir Ernest Marsden, whose experiments (with Geiger) led immediately to the nuclear hypothesis. The 450 delegates welcomed on the opening day by the Vice-Chancellor of the University (Professor W. Mansfield Cooper) were drawn from 32 countries and represented many different institutions. Particular aspects of Rutherford's work were underlined by the presence during part at least of the conference of Sir Charles Darwin, Sir James Chadwick and Sir John Cockcroft. In replying to the Vice-Chancellor the President referred briefly to his own association with Rutherford and tendered the gratitude of the delegates to the University for its hospitality. The meeting was then addressed briefly by the present Langworthy Professor (B. H. Flowers) on the facilities for communication between speakers and audience and with no further delay the conference entered upon its scientific business. This continued from Monday morning until Friday evening, with a break for an excursion to Chatsworth House on Wednesday and for a special commemorative session on the Tuesday afternoon, followed by a Congregation of the University. Appropriate Honorary Degrees were conferred upon Sir Ernest Marsden, V. F. Weisskopf, Aage Bohr and G. Racah, and upon Miss Annie Ellis, who worked as a young graduate in Rutherford's laboratory at the beginning of her career in the teaching profession.

The conference was appropriately devoted to nuclear physics rather than to elementary particle physics. To the possible disappointment of the Local and National Press, no new discovery of outstanding interest was announced; the addresses were mainly concerned with consolidation of, and corrections to, already established concepts. Rutherford, doubtless, would have found some of this tedious, since his was the heroic age of the subject, but he would have taken pleasure in the devotion and enthusiasm of the many young physicists who filled the conference hall. Each delegate received on arrival a small box file containing about 200 short contributed papers on a wide variety of subjects, but the main presentation of material was by the rapporteur method and on the whole individual contributions received only sporadic mention. At the conclusion of each group of invited talks by the rapporteurs a printed copy of each talk became available; these were in existence before the conference began and, together with the submitted papers and discussion, will later form the official record. The invited talks, with one or two notable exceptions, were heavily biased in the direction of theory and interpretation, but this was understandable, since a symposium on Nuclear Instruments was to follow (at Harwell) on 11th–12th September. The difficulties in nuclear physics indeed often arise in interpretation and the rapporteurs presented the delegates with a series of expert reviews which clarified many obscure points. Discussions tended to be short, but nevertheless the scientific impact of the meeting was considerable. In the following summary the official division of the available material into sessions is adopted, although it is not easy to draw sharp dividing lines in such a many-folded subject.

High energy investigations of nuclei (Chairman: J. Cassels)

Rutherford's experiments on nuclear transformation showed, as early as 1919, how the structure of nuclei might be studied, but by the time of his death in 1937 little more had emerged in this field than a realization that nuclear models must be based on a neutron–proton structure and that the forces between these constituents must be strong. The detailed properties of these forces can be inferred by a consideration of the regularities in the nuclear structures for which they are responsible, or, more directly, by a study of nucleon–nucleon collisions. It is also possible to predict some of the properties from theories of the type first proposed in 1935 by Yukawa, in which a π-meson is exchanged between two nucleons in much the same way that an electron is exchanged between two hydrogen atoms in the binding of the hydrogen molecular ion. It is now generally agreed that

it is useful to consider the force between nucleons as containing a long range attractive component, effective for internucleon distances of about 1·2f, and a shorter range repulsive component due to a 'hard core' extending over a distance of about 0·4f. H. P. Noyes, in an invited paper, reminded the conference that the theoretical treatment of the long range component of the force, as due to one-pion exchange in the sense of Yukawa, was promising but not yet complete. The theory of the hard core is far less satisfactory and may involve two or three pion exchange or strange particles. Only a phenomenological description of this short range force can at the moment be given. At a later stage of the conference Dr. Noyes described recent experiments at Berkeley on antiproton annihilation in a hydrogen bubble chamber which had given evidence of a sharp resonance between 3 pions at a total energy of 785 MeV.

The general properties of the nucleon–nucleon force which must be explained by any successful field theory are the saturation of binding energies of nuclei and scattering and polarization cross sections. It is already apparent from the experiments that a theory which includes dependence on spin and isotopic spin and which contains several varieties of non-central force is necessary. Elucidation of these independent contributions by experiments at low energies is hindered because the effects are small, and attention must be turned to higher energies, at which interaction in several states of orbital motion becomes possible. For the energy range up to 345 MeV the most extensive interpretation of experimental data in terms of one-pion exchange effects has been made by Breit and his co-workers at Yale University. Professor Breit made some comments on this work, which yields reasonable values for the pion–nucleon coupling constant. It also supports the idea that nuclear forces are charge independent, although it would not exclude a deviation from independence of the order of 3%, which was suggested in a later session on β decay. Dr. Noyes concluded with a survey of some of the theoretical attempts to describe the two-pion exchange process.

The theory of nuclear forces excited so much attention that R. E. Peierls led a further discussion on the subject in the last session of the conference, which had been reserved for topics of major interest. This led to general agreement that the nucleon–nucleon interaction must be described as far as possible in terms of two-body forces only, permitting velocity dependence, and that many body forces should not be introduced until some adequate guidance on how to do this was available from meson theory. The future of this subject is still a long and challenging one both on the experimental and theoretical fronts.

The contribution which fast particle scattering from *nuclei* can make to our knowledge of nuclear structure was surveyed in invited talks and discussion extending over two of the days of the meeting. In the first session H. McManus compared electron and nucleon scattering and emphasized that for particles of short wavelength high multipole nuclear transitions can be excited, since the long wavelength approximation of radiation theory no longer applies. The main types of information provided are (*a*) the nucleon distribution (from elastic scattering); (*b*) the quantum numbers of nuclear states, particularly those of collective motion (from inelastic scattering); and (*c*) the momentum and energy of nucleons in specific nuclear shells (from reactions such as p–2p). From the nucleon density distribution and from an assumed internucleon force it is possible to calculate an effective potential (optical model potential) in which the nucleons of a nucleus move. This is a convenient step in the model calculation of nuclear properties, which is an approach to structure problems that is more tractable, but less rigorous, than the solution of the many-body problem. In the case of inelastic nucleon scattering useful information can be obtained by measuring the polarization of the scattered nucleon, which may in part arise because of direct 'spin flip', of the particle. The experiments, made mainly by observing the angular correlation between scattered protons and gamma rays, indicate that this effect is small, at least for even–even nuclei. The p–2p reaction is particularly relevant as a check on theories of nuclear structure, as was first shown at Uppsala, since it determines the binding energy in the nucleon shells from which the second proton is extracted. This type of experiment has been continued with high resolution at Orsay. The effects of distorted waves in such experiments have been discussed, but they seem unable to account for an interesting anomaly in the results for ^{6}Li and ^{7}Li; here some form of cluster, e.g. (α + d for ^{6}Li), seems necessary. The predominantly theoretical aspect of this subject, which might well have excited Rutherford's impatience as the first morning wore on, was relieved in discussion by A. B. Clegg who described simple, Rutherford-type experiments made with the Harwell cyclotron. These very convincingly demonstrate the connection between inelastic scattering cross sections and the nuclear multipole moments derived from Coulomb excitation or lifetime measurements. This type of work, together with that on the p–2p reaction, underlines the usefulness of *high energy* (~150 MeV) investigations of complex nuclear structure, as distinct from the more familiar techniques of nuclear spectroscopy.

The final talk of the first session was by R. H. Dalitz on 'Hypernuclei and the hyperon–nucleon interaction'. It would be presumptuous here to attempt any summary of this expert account of an exciting new subject, but it may be noted if the hyperon–nucleon forces are known, then the problem of the binding of hypernuclei is just the usual problem of nuclear structure.

Collective motion in nuclei (Chairman: A. Bohr)

Although calculations of nuclear level properties based on independent particle models have had many outstanding successes, the collective aspect of nuclear motion is also evident in the behaviour of nuclei between closed shells. The forces between nucleons in nuclear matter are not necessarily those between nucleons in a free state, and in order to describe collective motion it is usual to think of a force of long range, comparable with a nuclear diameter, together with a force of short range, comparable with the internucleon distance. The former is responsible for deformations and for large-scale vibrations and oscillations, while the latter acts between like nucleons of opposite spin and orbital momentum, and is known as the pairing force. In an assembly of fermions, pairing forces lead to momentum states which are in essence just those discussed by Bardeen, Cooper and Schrieffer in their theory of superconductivity. There results an energy gap between the ground state of an even configuration of particles and the lowest excited state, in which a pair of correlated particles have changed their state of motion.

Another aspect of motion in a finite nucleus, perhaps more in the spirit of the shell model than of the collective model, is the coupling between an excited particle and the hole in the filled shell which it has left. The energy levels arising from this interaction were discussed by G. E. Brown as part

of a general account of the theory of collective oscillations. The predictions of this theory have been verified in impressive detail in the case of the odd parity states of the nucleus ^{16}O, in which a p particle transfers to the s or d shell. Experimental evidence for collective oscillations in general was reviewed by B. L. Cohen, who also showed how in a combination of shell model states one resultant state is always shifted down in energy and acquires a large collective multipole moment. A survey of the main experimental techniques for investigating collective states was given; these are essentially Coulomb excitation and inelastic particle scattering; the stripping reaction by contrast excites mainly single particle levels. The experimental evidence on the existence and location of 2^+, 3^- and 4^+ levels was summarized; the 1^- level of collective oscillation is well known as the giant resonance of photodisintegration. A picture of the excitation levels of a nucleus such as ^{58}Ni now begins to emerge; if examined under high resolution, the fine structure levels of the Bohr theory appear; if examined by the (d, p) reaction at low resolution, the reaction cross section is high in the region of single particle levels of a potential well, in which the absorbed nucleon is added just above the Fermi surface; if examined by the (pp′) or ($\alpha\alpha'$) reaction the levels of collective oscillation, in which there is a particular phase relation between some of the fine structure levels, are seen. The successes of the shell model calculation in predicting vibrational phenomena were, however, not allowed to divert the attention of the conference from the collective rotational states. H. E. Gove presented an impressive interpretation of the spectrum of ^{20}Ne in terms of rotational bands and Aage Bohr, in a good-humoured talk towards the end of the conference, emphasized the increasing knowledge of such bands in many deformed nuclei as a result of heavy ion excitation. The methods of introducing rotational properties into the formal structure of the shell model were discussed by S. Meshkov. More than fifty papers submitted for the proceedings testified to the healthy state of the collective model in nuclear spectroscopy.

The nuclear ground state (Chairman: R. E. Peierls)

When one considers the states of a finite nucleus rather than the binding energy of infinite nuclear matter, the size and shape of the nuclear surface become important nuclear parameters. In a talk which many will remember for its studied element of the picturesque, D. H. Wilkinson outlined the reasons for his faith in the existence of a small number of α particles as distinguishable entities in the nuclear skin. If any particles are to exist in such a situation there is little doubt that Rutherford would have preferred them to be α particles, but it may be questioned whether even he would have sanctioned their postulated role as elements in a recurrently inverting pin-cushion. The more solid grounds for this hypothesis are the frequent observation of α particles knocked out of nuclei by high energy nucleons, the anomalously high emission rates for certain naturally occurring α emitters and above all the phenomena of K^- meson absorption. In this process the tracks observed in nuclear emulsions and bubble chambers indicate a high probability that a K^- particle interacts with a correlated pair of nucleons. The interaction cross section is very large, and the process is confined more to the tail of the nucleon density distribution than in the case of the π-meson absorption. These ideas, which were rendered quantitatively, provoked a lively discussion in which at least one speaker felt that the evidence other than that from K^- studies was susceptible of other equally valid interpretations. Our main knowledge of the general outline of the nuclear surface (above which the α particles are supposed to extrude) comes from electron scattering experiments and from optical model analyses of heavy particle scattering data. A summary of many experiments of the latter type, given by P. E. Hodgson, showed how the parameters of a Saxon–Woods potential or a surface Gaussian potential could be derived consistently, given sufficient information, by parameter searches. It appears that weakly-bound bombarding particles such as ^3He are to be preferred for surface studies, presumably because they are unable to penetrate to the nuclear interior, but for this same reason such particles do not explore the inner nuclear potential. Values for the radius and diffuseness parameters of the potentials were given for several different incident particles. Similar analyses for a range of proton energies are being undertaken by Melkanoff, Nodvik and Saxon.

If the nucleon–nucleon force is known the binding energy in nuclear matter may be calculated using the powerful methods developed by K. A. Brueckner. Dr. Brueckner remarked during this session that such calculations had now been found to be more sensitive to the precise form of internucleon force than had been thought and that an interaction with a long range tail of the type predicted by meson theory should be used. He also gave an invited talk on the momentum distribution in a finite nucleus. The long-standing difficulty of reconciling the apparent existence of single particle orbits in a situation in which highly correlated motion is expected must now be interpreted in terms of the motion of 'quasi'-particles. These are nucleons with a meson cloud which can be shown, as in analogous fermion systems such as an electron gas or liquid ^3He, to fill single particle states. The technical details of current approaches to the many body problem were discussed in a talk by J. S. Bell. This attracted some attention because of the appearance of many body forces as a result of a somewhat special transformation contrived to describe the hard core.

Direct interactions (Chairman: H. H. Barschall)

The processes discussed by McManus in the first session are closely related to a large body of nuclear reaction phenomena usually classed as *direct interactions*. These differ from the compound nucleus type of reaction because the primary process is an interaction between the incident particle and a nucleon in the nuclear surface. The fifth session of the conference (Thursday morning) was devoted to these processes, of which the deuteron stripping reaction is the best known and most exhaustively studied. The plane wave theory of this reaction, first given by Butler, has permitted the deduction of the parity change between levels of nuclei throughout the Periodic system, and has given important confirmation to the predictions of the single particle shell model. Moreover, many different types of reaction have been found to proceed by an essentially similar mechanism for sufficiently high energies ($E \sim 10$ mev or more) and to offer similar information. In clear and informative reviews of this field, including the papers submitted to the conference, J. B. French and H. E. Gove discussed respectively the theory of direct interactions and the experimental evidence for the phenomenon, but since theory and experiment are here inextricably linked, these talks overlapped to some extent. Emphasis was placed on the fact that the simplest plane wave theory of the process had on the whole given an excellent account of many angular distributions and of relative cross sections. Distorted wave calculations based on realistic

potentials between nucleons and nuclei can now be made relatively easily and certainly give a better account of stripping phenomena in general, but it is interesting to enquire why the plane wave theory works so well. In a comment following the invited papers N. Austern suggested that this might be understood since the range of energies covered by the experiments is really not very great, and the interaction radius is determined by curve fitting, but it seems likely that the complete explanation has not yet been offered. The theory of direct interactions involving transfer of two or more particles to (or from) the target nucleus has as yet not advanced beyond the plane wave stage, but much experimental data on reactions such as (^{3}He, p) and (^{3}He, n) is now available for analysis. Some of these data appear to indicate the occurrence of 'heavy particle' stripping, which is one reason for the appearance of backward peaks in an angular distribution.

Inelastic scattering of α particles has received much attention in the last few years and the theory of this process given by Blair, in which scattering takes place from an effectively 'frozen' deformed nucleus, was exemplified and discussed. It is possible in this theory to avoid the inconsistency between plane wave treatments and the known strong nuclear absorption of α particles. The technique has led to the identification of an important series of 3^- levels in even–even nuclei.

It was felt by the invited speakers and by several of those who contributed to the discussion, that if suitable corrections for scattering of the incident and emergent particles were included in the theory then excellent agreement with experimental angular distributions could be expected. Ideally the potentials to be used in making these scattering corrections should be determined directly by experiments on the unbound systems concerned, e.g. by scattering protons from the residual nucleus in the case of the (d, p) reaction. Several examples of this type of calculation using either observed or assumed scattering potentials were presented; the ability of the theory to account for the angular distribution in the case ^{206}Pb(d, p)^{207}Pb for a range of deuteron energies from 8 to 15 MeV is particularly impressive, and gives confidence that the distorted wave treatment permits this important technique of nuclear spectroscopy to be applied throughout the Periodic system. In the discussion W. Tobocman illustrated the same point convincingly for light nuclei by discussing the correlation of data on the ^{3}He(d, p)^{4}He, ^{3}He(d, d) and ^{4}He(p, p) processes.

One of the consequences of the distortion of incident and emergent waves in the stripping process is polarization of the outgoing nucleon. L. J. B. Goldfarb reviewed progress in this field since the Basel conference on polarization phenomena and noted the availability of accelerated beams of polarized nucleons and deuterons in an increasing number of laboratories. The sign of the polarization of the stripped protons depends on the spin of the nuclear state concerned and on the type of distortion present; it appears that deuteron wave distortion is more important than proton distortion in most of the cases so far studied. In the special case of stripping with $l = 0$, polarization can be due only to spin-dependent forces, which are not usually important in this phenomenon, although such effects must be present if the observed polarization exceeds $33\frac{1}{3}\%$. Observations of the spin-dependent type of polarization then determine the spin-orbit part of the optical model potential for deuterons interacting with nuclei. Reciprocal relations between polarization in direct interactions and asymmetries observed with polarized beams can also be established. In the discussion following this paper, S. T. Butler offered an attractively simple picture of the origin of polarization in a stripping reaction. Essentially the observed effect comes from two 'sides' of the nucleus defined by the direction of transferred linear momentum, but the contribution from one side is attenuated by nuclear absorption more than that from the other, and a predominance of one sign of spin direction then results. For small angles the sign of polarization determines j value; for large angles spin dependent effects may be seen. The popularity of the investigation of direct interactions with the present generation of medium energy accelerators was shown by the large number (64) of papers submitted for this session.

Weak interactions (Chairman: M. Goldhaber)

In opening the Friday session the chairman reminded the meeting that although Rutherford often expressed a strong personal preference for α particle physics, he nevertheless collaborated with Robinson, Chadwick and Ellis to lay the foundations of quantitative β and γ spectroscopy. Like many other fields in which Rutherford was a pioneer, that of β decay has enormously expanded and we have today a good understanding of the mechanism of the β decay interaction (as distinct from the nuclear matrix elements which enter into transition probabilities). The talk by R. J. Blin-Stoyle consequently dealt with 'small effects' rather than with broad features and he selected for especial attention in the first place the difference between the theoretical and observed decay rates of the μ-meson. If there is a universal Fermi-type interaction coupling electrons, nucleons, muons and neutrinos, then the discrepancy of 2% is significant but it may be removed by postulating a charge dependence of 2–3% in the intrinsic nuclear forces. Such a dependence is perhaps also indicated by the fact that isotopic spin selection rules for nuclear reactions do not appear to be absolute. An alternative explanation of the muon discrepancy is that the flow of charge in the β decay process, which involves nucleons and therefore the π meson field, is not conserved. This somewhat theoretical proposition can be tested experimentally by careful measurements of the shape of the β spectrum of even isobars such as ^{12}B and ^{12}N, and also by looking at the angular correlation between β particles and α particles in the decay of ^{8}Li. W. A. Fowler later reported on measurements of the β spectra which indicate that the vector current is conserved; he also announced an extremely accurate determination of the comparative half-life for the important Fermi-allowed transition in the ^{14}O decay. In continuation Dr. Blin-Stoyle pointed out small deviations of the decay data for certain simple mirror nuclei from expectation. These appear to be associated with inexact evaluation of the nuclear matrix element for the Gamow–Teller interaction. The β decay data can be used to correct this matrix element, and the nuclear magnetic moments can then be deduced. When these are suitably adjusted by relativistic and wave function corrections, there is still a residual discrepancy between experimental and computed magnetic moments which is ascribed to a mesonic exchange term.

The weak interaction session also included a report by C. Rubbia on muon capture in nuclei, in which results obtained by Hildebrand on the important capture by hydrogen were described. The capture rate is consistent with the known β mechanism and gives further support to the idea of a universal Fermi interaction by a route independent of uncertainties connected with nuclear matrix elements. In another invited talk S. Moszkowski showed how, given a

complete understanding of the β-interaction, it becomes possible to obtain the nuclear matrix elements reliably and thus to test nuclear models. The decay of deformed nuclei, the possible relaxation of isotopic spin conservation and the spectrum of RaE were cited as topics for special consideration. Towards the end of the session V. L. Telegdi described an experiment (as yet incomplete) in which the hyperfine structure of a μ mesic atom is to be examined by detecting the products of decay from the two hyperfine structure levels. This will determine the spin dependence of the muon–nucleon capture process.

Limitations and needs in instrumentation for nuclear physics (Chairman: J. Mattauch)

In an afternoon session devoted to two specially invited papers on quite different topics, S. Devons, a former student in Rutherford's laboratory at Cambridge, and one of his successors in the Langworthy Chair, discussed nuclear instruments, with particular emphasis on studies of the electromagnetic properties of nuclei. This stimulating talk well illustrated the connection between the development of techniques and the reward of this effort in the extension of nuclear information. The area in which most extension is now required is perhaps that of the measurement of electromagnetic properties of excited states. Here the limitation that arises is due to the lack of suitable magnetic fields for causing spin precession and observable rotation of an angular correlation pattern. For states with lifetimes of 10^{-12} sec or less the field at the nucleus resulting from the presence of a single K-electron could in principle be used; even larger fields would be available if a μ-mesic atom could be employed. Another familiar limitation in electromagnetic studies is the absence of a source of monochromatic γ radiation of continuously variable energy. This is being overcome by the use of the radiation from fast positrons undergoing single quantum annihilation in flight and the first results of photonuclear experiments with this radiation are now appearing. Other limitations discussed related to the topics of β decay and nuclear collisions for which a general survey of the technical developments of the last few years was offered. In conclusion the speaker referred to Rutherford's use of the 'high energy' approach, in that he resolved the structure of the atom by the use of α particles of energy 8 MeV. An extrapolation of this attitude leads us directly to the use of the high available energies of the present day in probing nuclear structure. It might be well, however, if we were to exercise some of Rutherford's inspired caution in undertaking some of the more elaborate and expensive investigations.

Rutherford and nuclear cosmochronology (Chairman: J. Mattauch)

This paper, which deservedly seemed to receive the greatest acclamation of the whole conference, and which would certainly have gladdened Rutherford's heart, was given by W. A. Fowler of the California Institute of Technology. In a fluent, well-organized and elegantly documented account of the origins of the elements, the speaker paid due tribute to Rutherford's pioneer work at Manchester in determining the half-lives of the naturally occurring radioactive substances. His subsequent suggestion of the helium and lead methods for finding the age of terrestrial minerals led, even during the lifetime of Lord Kelvin, to a revision of the age of the earth from about 10^8 years to more than 10^9 years. A surprisingly accurate estimate of the age of the elements was also made from the lifetimes and abundances of the uranium isotopes. Rutherford also anticipated, in broad outline, the modern theory of the generation of solar energy by transmutations, even before he had come to the hypothesis of the nucleus. The recent astrophysical work of Burbidge, Burbidge, Fowler and Hoyle has not essentially altered the geological and radioactive estimates of the age of the solar system, which stands at $4{\cdot}5 \times 10^9$ years. Astronomical research during the last 10 years has, however, pushed the age of the universe up to 5–20 $\times$ 10^9 years if an act of creation is assumed and to infinity if continuous creation takes place. The contribution of the Californian workers has been to calculate the processes of nucleosynthesis using terrestrially measured nuclear cross sections and to set the age of the Galaxy in which these processes have taken place at $20 \pm 4 \times 10^9$ years. Dr. Fowler elaborated the steps leading to these conclusions with admirable clarity. The age of the universe is determined from the red shift in the spectral lines of receding galaxies once the scale of astronomical distances has been established, and it is for this latter magnitude that more reliable values have recently become available; despite the increase in reliability a clear-cut decision between exploding theories of creation and the steady state hypothesis is not yet available. Neither theory, of course, accounts for the genesis of the primordial hydrogen. The age of the stars in the Galaxy is obtained by assuming that a star becomes a red giant (it can be identified astronomically as such from the Hertzsprung–Russell diagram) when its hydrogen has been converted to helium. The rate for this process is calculated from nuclear data and the age of those stars which have 'recently' exhausted their hydrogen is then obtained. The results however are not so reliable as to make an independent estimate of the age of the elements from nuclear data unnecessary. Dr. Fowler traced the details of this estimate, in which it is assumed that elements up to iron are formed by charged particle reactions, and heavier materials by neutron capture. The nuclei ^{238}U and ^{235}U are supposed to have been formed only in conditions existing in supernovae and to have been injected into the Galaxy from these stars. The combination of these processes indicates a Galactic age consistent with the astronomical estimate based on the hydrogen-burning time. This paper presented an impressive picture of the growth of our knowledge of the physical universe and of the vital part played in stellar evolution by nuclear reactions.

By the end of the fifth day of the conference everybody was conscious that a great deal had been said and that it should be possible to draw some conclusions about the present state of nuclear physics. This considerable task was courageously undertaken by D. R. Inglis whose conference summary perhaps rather emphasized the theoretical contributions, but showed how the links between mesons and forces, between forces and models and between models and actual nuclei were steadily becoming stronger. The excitement of speculation and experiment in academic nuclear physics which is a direct outcome of Rutherford's work should not blind us, the speaker remarked, to the fact that because of that same work and its consequences, we live in precarious times. On this sober note, and after a short speech of thanks by P. Huber on behalf of The International Union of Pure and Applied Physics, the assembly dispersed.

Department of Physics,
The University,
Edgbaston,
Birmingham 15.

W. E. BURCHAM
19*th September* 1961

Proceedings of the Symposium on Tunnel Diodes—London, February 1961

Abstract

The meeting covered the theory, physical characteristics and circuit applications of tunnel diodes, and the proceedings constitute a summary of the eight papers delivered.

THE Symposium was jointly organized by The Institute of Physics and The Physical Society (Electronics Group) and the British Institute of Radio Engineers. At the time of the announcement of the Esaki (tunnel) diode in 1958, an advanced stage of technology had been reached in silicon and germanium and a number of III–V compound semiconductors were available for technical exploitation. In addition, there existed considerable information on the band structure of these materials, and a general understanding of the physics associated with p–n junctions. The necessary groundwork therefore existed for the rapid development of tunnel diodes as high-speed switches or as high frequency generators in the kilomegacycle range. The expanding field of computers, with a continued pressure for increased speed of operation, has enhanced this rate of development still further.

The tunnel diode is essentially a narrow p–n junction device in which the addition of doping elements to the parent semiconductor crystal lattice has been pressed to the limit, resulting in a barrier width of less than 100 ångstroms, and majority carrier densities $\sim 10^{19}$ per cubic centimetre. The junction width is in consequence comparable with the mean spacing between impurity atoms. Quantum mechanical theory indicates that there will exist a finite probability of electron tunnelling through the potential barrier of the junction provided that conservation of momentum occurs, and that empty energy levels exist to which the electrons can tunnel.

In the forward direction the tunnelling probability is at first enhanced by the application of forward volts, causing a rise in diode current to a characteristic peak value. A region of negative conductance is then encountered, the current falling with further increase in applied voltage, to a second characteristic minimum or valley figure. Tunnelling ceases when the application of further forward voltage causes the bottom of the conduction band to fail to overlap with the top of the valence band in the conventional energy diagram. A further increase in forward voltage causes normal diode behaviour, with injection of minority carriers.

With reverse bias, electron tunnelling probability in the opposite direction is increased, and the diode appears to exhibit zero reverse breakdown voltage.

The peak voltage depends, not on energy gap, but on Fermi-level, whereas the peak current shows a critical dependence on effective mass and impurity density. The effective mass is frequently uncertain, and may be a small fraction of the true electron mass in certain materials.

The valley current should reach zero as the tunnelling probability goes to zero, but an excess current is normally observed, attributed to electrons tunnelling to or from states within the forbidden energy gap. The density of such states may be subject to technological control, if due to crystal imperfection or extraneous impurity, or may increase with time if due to radiation or bombardment damage.

Capacitance of the tunnel diode, due to the extremely narrow barrier width, may reach $\sim 1\ \mu F/cm^2$. For an effective high speed device, the cross-sectional area will be extremely small, with consequent problems of series inductance due to the package. The tunnelling phenomenon has a time constant $\sim 10^{-12}$ second, so that provided minority carrier injection is avoided, fast switching times ($\sim 10^{-9}$ second) can be achieved with existing device parameters. The small cross-sectional area, linked to limited voltage excursions, will, however, place limitations on the power handling capacity.

Materials widely used for tunnel diodes have included germanium, silicon, gallium arsenide and indium antimonide. In gallium arsenide, transitions do not involve a change in momentum; with germanium and silicon, however, conservation of momentum requires the tunnelling process to be phonon-assisted by the absorption or emission of a phonon as an intermediate step.

It follows that the tunnel diode has a dual role to play, firstly as a high speed switch in fields of application where transistor performance is inadequate and secondly as a scientific tool for the examination of band structure of semiconductors in a manner more direct than was previously possible.

It is evident that tunnel diode fabrication places a premium on the ability to dope semiconductor materials heavily, and is causing a new emphasis to be placed on methods of crystal growth which permit introduction of abnormally high impurity concentrations without introducing structural imperfections or localized inhomogeneity of impurity.

The material requirements for a effective tunnel diode may here be restated.

(i) The Fermi-level on both sides of the junction must be inside the conduction and valence bands.
(ii) If both band-gap and effective mass are small, high impurity concentrations are required.
(iii) The 'normal' diode current due to injection of minority carriers must be low.

The requirement for high impurity concentration if attained by the use of high process temperatures, may involve not merely a solubility condition for the intentional doping agent, but also the avoidance of stray impurity diffusion.

With indium antimonide, the band gap is of the order 0·2 ev, and the effective mass $m_r \sim 6{\cdot}5 \times 10^{-3}$. In consequence, an impurity concentration of $2 - 6 \times 10^{17}$ atoms/cm^3 is necessary, which is feasible with tellurium doped n-type material. P-type alloy regions of adequate impurity concentration can be fabricated using cadmium or indium–cadmium.

The third requirement, that the 'normal' diode current be small, necessitates cooling the devices to $-77°$ C. At this temperature, peak to valley current ratios of twenty are

observed, with the peak and valley currents occurring at 25 mv and 100 mv respectively.

Indium antimonide tunnel diodes should possess considerable advantages over germanium devices, by virtue of the lower doping levels required, leading to a lower capacitance per unit area by a factor four. In practice this has not been clearly verified. The noise figure should also be lower, because the voltage excursions are smaller than in the case of germanium.

A real fabrication difficulty is the provision of a suitable etch which will remove InSb without removing the alloy dot material. On balance it appears that indium antimonide diodes are unlikely to find widespread practical application.

It is in the investigation of band structure that indium antimonide tunnel diodes have been proved useful. A less complicated band structure than germanium, allied to a sensitivity of tunnelling to magnetic field, provide a useful area of research.

Whilst the magnetic effects suggest a controlled device, the reduction in tunnelling is only a few per cent at practicable fields, although at 150 000 gauss the peak current may be reduced by a factor of twenty.

A considerable volume of pioneer work on gallium arsenide has been carried out at the Services Electronics Research Laboratory, Baldock, and techniques used for tunnel diode fabrication were described to the conference. Gallium arsenide is first zinc doped at 1000° C in a closed system to produce p-type material with an impurity concentration of 2×10^{20} atoms per cubic centimetre. N-type regions are made by alloying tin dots, and a caustic electrolytic etch is used to finish the device to size. The 30 mil dice is mounted in a pill structure, resin filled.

Whilst peak to valley current ratios of 60 : 1 can be observed, a more normal figure is 15–20, with peak current densities of 16 000 A/cm^2. Under pulse conditions this value may be extended to 150 000 A/cm^2. It is found that the peak current is reduced if the alloying temperature is raised from 475° C towards 800° C. The figure of merit, in terms of the ratio peak current to capacitance, is quoted as 13 mA/pF. Parameter measurement raises some difficulties since the device series resistance is less than one ohm, and series inductance is 4×10^{-10} henries. The equivalent resistance must be considered as consisting of a series inductance and resistance with a negative conductance shunted by the diode capacitance. Under measurement conditions, the diode is shunted by a sputtered non-inductive resistor for stabilization, and measurement carried out on a transfer-function bridge.

Diodes of this type have been employed as oscillators in both strip-line and cavity configurations and have delivered 5 mw at 500 mc/s; operation at 4·6 Gc/s has also been achieved.

Gallium arsenide offers a wider range of pseudo-linear negative conductance on the voltage axis than does germanium, and low series resistance is readily achieved. The high temperature performance is better than germanium. The valley current is found to increase monotonically with increasing temperature, and the peak current varies slowly, the ratio being ~3 : 1 beyond 300° C. The peak current falls by ~20% between room temperature and liquid nitrogen.

The conference discussed in some detail the ageing effects of gallium arsenide diodes, opinion being divided on its inevitability. It was suggested that the fall in peak–valley ratios observed in service was due to junction non-uniformity, resulting in the development of local hot spots. In these high temperature regions, impurity diffusion could occur causing serious modification of junction characteristics. A limitation of peak current would tend to limit the effect. As the diodes are heavily doped, it was generally agreed that the ageing effect could not be associated with the surface.

Further discussion centred on the stability of tunnel diodes to radiation and bombardment damage and American data was quoted to support the view that fast neutrons cause permanent damage, increasing the valley current and decreasing the voltage swing available. Germanium, silicon and gallium arsenide all show noticeable changes in fast neutron doses of $10^{15} - 10^{16}$ nvt. By 10^{17} nvt, germanium and silicon diodes have become inoperable, while gallium arsenide still shows appreciable negative resistance.

The design and manufacture of both tunnel diodes and unitunnel (backward) diodes were then discussed using germanium as base material. The established technology of the material, its ready availability at an economic price, and the degree of control possible, may outweigh for mass usage the relative disadvantages of germanium tunnel diodes. Peak and valley currents occur at 50 mv and 250–400 mv respectively, and this voltage separation is found to shrink with increasing temperature.

Values of peak current are quoted to 10% or less, and the fabrication technique uses a monitored electrolytic etch giving finish tolerances of ±2%. Gallium doped germanium is used with impurity concentrations of $1–7 \times 10^{19}$ atoms/cm^3 for the base material (p-type) with antimony or gold–antimony alloy dots, six mils in diameter, to provide the n-type regions. Some improvement in performance is achieved by addition of arsenic to the alloy, to counter the low segregation coefficient of antimony in the regrowth area of the junction, thus maintaining a high value of impurity density in this region.

Whilst the tolerance of peak current is less than ten per cent, it emerged in discussion that the capacitance cannot be held to this limit.

Three alternative packages were described, ranging from conventional header assemblies with two wire lead-in, with series inductance values of 8 nH, 'top-hat' packages where this value had been reduced to 4 nH, and a symmetrical pill-type assembly, where the inductance was less than 1 nH.

The circuit engineer is requiring higher figures of merit in terms of milliamperes per picofarad, and it is possible that these requirements may be met in the future by the use of epitaxial techniques. Epitaxial methods may permit the production of lower series resistance diodes, and introduce economies in fabrication.

In discussion, the temperature stability of the germanium tunnel diode was queried with a view to establishing whether ageing effects (analogous to the gallium arsenide case) were experienced. It was stated that after one hour at 400° C, the peak current was halved, believed due to local arsenic diffusion in the n-type regrowth region. It was emphasized that the temperature behaviour of the germanium tunnel diode, by virtue of being constructed from heavily doped extrinsic or degenerate material, is better than the germanium transistor with which it is likely to be associated in computer circuits. Hence the transistor is the component most likely to set the environmental limits which will be permissible in a given application.

The unitunnel, or backward diode, was described as an extension of the tunnel phenomenon, and is made by similar techniques to those used for the normal tunnel diode. The n-type doping level is, however, reduced to ~10^{18} atoms/cm^3.

Under these conditions the forward current peak, corresponding to a maximum tunnelling probability, disappears, whilst the reverse breakdown due to tunnelling, commencing at zero volts, gives a fast rise in current as the reverse volts are increased. Rectification is thus effectively inverted, compared with a conventional semiconductor diode, the forward direction of the diode becoming the high impedance condition, over a restricted range. Rectification ratios of one thousand at ± 200 mv have been obtained, making the backward diode suitable for use as a coupling element in tunnel diode applications. It was recognized, however, that although the backward diode operates at high impedance in the nominal forward direction, the junction capacitance is correspondingly high due to the reduction in space-charge region width by the applied voltage.

The addition of a third control electrode to the tunnel diode structure would enhance its flexibility of circuit application, and considerable thought has been given to the achievement of external triggering. The small physical size of the tunnel diode does not lend itself to the introduction of a physical electrode or an auxiliary layer, and this fact, coupled with the insensitivity of the device to light and magnetic fields, has led to the development of a double-based tunnel diode, and the exploration of its characteristics.

In the structure examined, a germanium dice 30×5 mils is utilized, and two ohmic alloy contacts are made to the degenerate base region in addition to the formation of a normal tunnel diode alloy structure on the upper side of the dice. It is found that the shape and position of the negative conductance characteristic can be materially influenced by the magnitude of the lateral current flow between the ohmic contacts. Since the tunnel characteristics are not readily represented by analytic functions the device performance was evaluated with the aid of a computer, and the trends established support the experimental results obtained on fabricated devices. No circuit implementation of the device was discussed, but the sensitivity of the control characteristic is believed sufficiently promising to warrant further investigation.

The conference turned, in its second session, from the consideration of the principles and techniques of fabrication of tunnel diodes, to their applications, principally in computers. A wide range of circuit functions was discussed and it was demonstrated that the tunnel diode, being a two-state device, can be utilized as both a logic and a memory element in high-speed systems. It can, in principle, replace both transistor and ferrite core to form a unified system, provided that cost figures and tolerances permit.

Due to the lack of current gain, however, and the small voltage excursions available, it seems probable that mixed systems of transistors and tunnel diodes will persist for some time. Since the tunnel diode, relative to the transistor, is insensitive to temperature and radiation, no new environmental problems in such systems are visualized.

Majority-logic systems have tended to be resistively coupled, and in consequence slow with poor tolerancing. To improve the tolerance position, operation in the voltage mode is recommended with tunnel diodes, which call in addition for circuit complication due to the need to make the logic system unilateral.

Isolation between input and output circuits can be achieved conventionally by insertion of a rectifying diode in each input path to the logic element. In tunnel diode systems the voltage excursions are low, and normal computer diodes may, in consequence, fail to give the required degree of isolation. The 'backward' diode (i.e. a lightly doped tunnel diode, resulting in the inversion of the normal rectification characteristic) offers a possible solution of this problem, which has yet to be fully explored, since it exhibits an adequate rectification and impedance ratio over the voltage ranges commonly encountered in tunnel diode circuits, whilst retaining speed associated with majority carrier operation.

The simplest utilization of the tunnel diode is shown in Fig. 1. If the supply voltage is less than the valley voltage

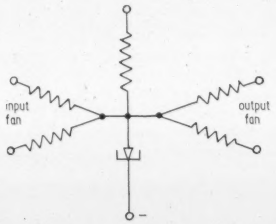


Fig. 1. Tunnel diode bi-stable circuit.

of the diode, the circuit will behave as a monostable multivibrator, and can be utilized as (i) a regenerative trigger circuit, (ii) a non-linear pulse amplifier, (iii) a generator of relatively uniform output pulses.

The load line of the tunnel diode will operate from a point on the IV characteristic just below the peak current to some chosen point on the valley portion of the characteristic.

Fast switching will be obtained, utilizing only majority carrier operation, with a load line having a negative slope slightly less than the negative slope of the diode. Such operation requires extremely stable bias voltage, and a limited number of elements on the input and output fan.

Slower switching permits a limited amount of minority carrier injection, and utilizes a more nearly horizontal load line. To allow for diode, resistance and supply voltage tolerances, the low impedance condition of the diode must be set well below the peak current point to permit non-critical operation.

The four basic transfer systems, using tunnel diodes, are shown in Fig. 2, with rectifying diode isolation.

The logic functions of 'inhibit' and 'inversion' are shown in their tunnel diode form in Fig. 3, where an input at X inhibits an input at Y. For a signal of 'ONE' at Y, and X at 'ZERO', the TRANSFER-IN clock sets B_2 to its high state and hence results in a negative output. If the input X is at 'ONE', the diode B_1 is set in the high state, and prevents the -250 mv level at Y from setting B_2; a positive output (a ZERO) will appear when the TRANSFER-OUT pulse is applied to D_1.

The function of memory with read-transistor is shown in Fig. 4. Reading is achieved through voltage mode selection using a conventional computer diode with a forward conducting potential of 250 mv compared with 500 mv for the tunnel diode.

Two cases arise:

(i) The tunnel diode is in the low voltage condition. By placing a 250 mv pulse on the word line, the diode D_1 will conduct at 300 mv (250 mv + 50 mv) and the tunnel diode will switch to the high voltage state.

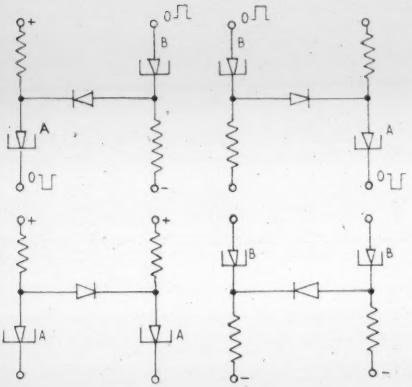


Fig. 2. Basic tunnel diode transfer systems.

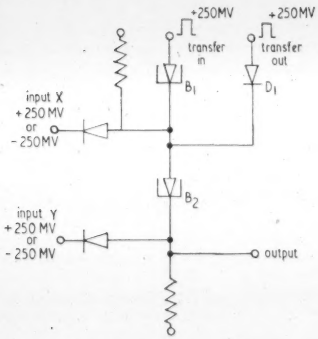


Fig. 3. Inhibit and inversion circuit.

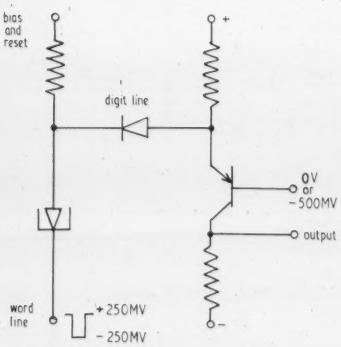


Fig. 4. Memory with read-transistor circuit.

(ii) Tunnel diode in the high voltage condition. No current would flow in the read line until the word line reached −750 mv (−500 mv −250 mv). Hence the state of the diode can be read by a negative pulse of magnitude between −300 mv and −750 mv, the tunnel diode being left in the high voltage state. New information can be written into the tunnel diode by first returning it to the low voltage condition and applying a pulse to the word line. The digit line is held at ground potential for a high voltage state requirement, or negative if a low voltage state is required.

Using this type of circuit a store utilizing the minimum number of components has been constructed, containing 8 digits of 8 words; 64 bits. Each read transistor has as many diodes connected to its emitter as there are words, with the sum of the diode capacitances acting as a limit to the speed of the store. Total access time is 40 ns with a read-out time of 10 ns.

With the increasing complexity of electronic equipment, where both complexity and speed considerations are demanding a reduction in the physical size, 'solid state' philosophies are being applied to computer needs. The elements used in large quantities in computers must of necessity be relatively cheap and the extension of 'solid circuits' to large numbers of high grade transistors would be uneconomic in terms of yield, although two-transistor elements have been shown to be feasible. The chosen circuit element should be small, robust, cheap, immune to surface effects, of low power consumption to permit high packing density and amenable to mass production.

It was suggested that the tunnel diode basically meets these requirements. The various application papers demonstrated that the device is sufficiently flexible to perform most of the computer functions hitherto carried out by transistors, and that a reduction in the number of components, always an aid to tolerancing and reliability, could be achieved.

Rather than pursue a philosophy of total integration, the view was expressed that 'packaging' of associated components, e.g. pairs of tunnel diodes, or associated tunnel diodes and resistors in shift registers, could usefully be undertaken, these functional sub-units being associated with printed or evaporated wiring networks.

In the concluding paper of the conference consideration was given to gain factors in tunnel diodes, in relation to

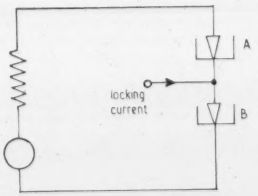


Fig. 5. Tunnel diode pair.

operation tolerances. It was concluded that energy gain could be high (~150), and is in fact higher than either power or current gain.

For circuit design purposes the tunnel diode characteristic must be formalized in a relatively simple manner, and it was suggested that the I–V characteristic could be satisfactorily represented by four pseudo-linear regions, characterized by four values of conductance, representing the rise and fall of peak current, the valley region, and the 'normal' diode current.

Each of these regions has associated with the rising voltage an increasing capacitance value, so that a switching time-constant proportional to the product of reciprocal conductance and capacitance may be allotted to each region for

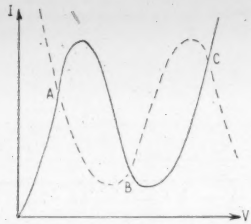


Fig. 6. Composite tunnel diode pair I–V characteristic.

design purposes. This formalization of diode behaviour is adequate only for trigger currents less than the peak diode current.

A basic combination for computer logic is the tunnel diode pair, and a wide range of circuit function can be built up. In these combinations, one tunnel diode acts as the load for the other. The behaviour of such a circuit can be examined by superposition of the two diode I–V characteristics, treating one as a non-linear load line representation. It will be seen that there are three possible points of intersection, A, B, C, of which only A and C are stable. The choice of state, A or C, can be made by use of an appropriate locking current applied at the junction of the two diodes, a locking current in giving a high voltage mode to diode B, and conversely a locking current out giving a low voltage state.

Tunnel diodes, cascaded, can be used for simple scaling, with automatic reset when combined with transistors, and the simplest form is shown in Fig. 7. Such a chain will trigger sequentially if the peak current ratings of the diodes are graded sequentially. Successful operation with up to five tunnel diodes in the chain has been achieved.

A number of other circuits including divider, phase splitter, and two phase shift register were discussed, which served to illustrate the wide circuit design flexibility.

In conclusion, it should be noted that the primary object of this joint meeting was not to present advanced research information, but rather to establish the 'state of the art' and induce a degree of mutual understanding between physicists

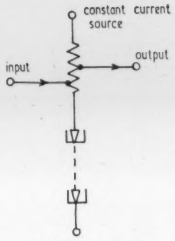


Fig. 7. Tunnel diode scaling circuit.

and engineers working in different areas of the subject. In spite of a wide divergence of emphasis in the content of the papers presented, this integration function was achieved, and the meeting has set a useful precedent.

Marconi Wireless Telegraph Co.,
West Hanningfield Road,
Great Baddow,
Essex.

I. G. CRESSELL
24*th October* 1961

Papers read at the Symposium

The tunnel diode
A. K. Jonscher, General Electric Co. Ltd., Wembley.

Indium antimonide tunnel diodes
K. F. Hulme, Royal Radar Establishment, Great Malvern.

Gallium arsenide tunnel diodes
J. J. Low and R. J. Sherwell, Services Electronic Research Laboratory, Baldock.

Theory, design and manufacture of tunnel diodes
G. N. Roberts, Standard Telephones and Cables Ltd., Footscray.

The tunnel triode
S. Amer and W. Fulop, Standard Telephones and Cables Ltd., Footscray, and Standard Telephone Laboratories Ltd., Harlow.

Tunnel diode applications
G. B. B. Chaplin, The Plessey Co. Ltd.

Tunnel diodes as 'solid state' elements
A. E. Brewster, Standard Telephones and Cables Ltd., Harlow.

Tunnel diodes as switching elements
R. W. A. Scarr, Standard Telephones and Cables Ltd., Footscray.

Indium antimonide tunnel diodes

by K. F. HULME, Ministry of Aviation, Royal Radar Establishment, Great Malvern, Worcs.

MS. received 16*th August* 1961; *paper presented at the Symposium on Tunnel Diodes on* 7*th February* 1961

Abstract

The small forbidden energy gap and effective masses in InSb make it of interest because the high tunnelling current densities needed for an electronically fast device can be achieved very readily with low doping concentrations. The major disadvantage from the viewpoint of electronic applications would be the need to cool the diode. The ways in which the study of InSb tunnel diodes has contributed to the understanding of the physical processes involved in tunnelling, are described.

INDIUM antimonide belongs to the class of semiconductors discovered by Welker (1952) which are compounds of elements from Groups III and V of the periodic table. Before considering tunnel diodes of InSb, it is appropriate to recall its most important parameters as a semiconductor.

The forbidden energy gap E_g between valence and conduction bands is small (0·18 ev at 300° K and 0·23 ev at 77° K); the effective mass of conduction electrons m_e is also small, being only 0·014 of the free electron mass m_0; the effective mass of the heavy holes m_{hh} is 0·18 m_0, and of the light holes m_{hl} about 0·014 m_0. The minimum of the conduction band and the maxima of the heavy- and light-hole valence bands (which are degenerate at $k = 0$) are all at the centre of the Brillouin zone. The electron mobility is high (77 000 cm^2 v^{-1} s^{-1} in pure material at room temperature and 600 000 cm^2 v^{-1} s^{-1} at 77° K). So far its applications have been in detectors of infra-red radiation (the small energy gap makes it sensitive out to 6–8 μ) and in Hall effect devices (here the high electron mobility is an advantage). It is only recently that serious efforts have been made to exploit the useful properties of p–n junction devices made of InSb. Because of its high electron mobility it has advantages as a material for fast diodes and transistors (Henneke (to be published) has recently described a fast InSb transistor). There is, unfortunately, one drawback to p–n junction devices made of InSb: they must be cooled well below room temperature (e.g. to 77° K) if the p–n junctions are to present appreciable electrical impedance. This is fundamental, and is a consequence of the small energy gap; there are too many thermally excited carriers at room temperature, and the height of the barrier which can be provided in a p–n junction is too small in comparison with thermal energies. The requirement of cooling also applies to tunnel diodes.

From Esaki's description* of how quantum mechanical tunnelling gives rise to a negative resistance region in the voltage–current characteristic of a sharp junction between heavily doped p- and n-type Ge, one would at first conclude that Esaki or tunnel diodes could be made from any semiconductor. Second thoughts would prompt the realization that the solubilities of the donor and acceptor impurities required to produce the heavily doped regions will be finite, and that it might be difficult to obtain an electronically fast tunnel diode with a high energy gap semiconductor. The circuit speed of a tunnel diode is not limited by any transit time effects, but by the junction capacity (C per unit area) divided by the value of the negative conductance due to tunnelling (G per unit area). Now C is a relatively slowly varying function of the doping concentrations on either side of the junction (P and N); in fact in rationalized units

$$C = \left\{ \kappa\kappa_0 e \Big/ 2\phi_0 \left(\frac{1}{N} + \frac{1}{P} \right) \right\}^{1/2},$$

where κ is the relative dielectric constant, e the electronic charge, and ϕ_0 the barrier height (roughly E_g for small bias). On the other hand, G will be approximately proportional to the tunnelling probability p of an electron impinging on the junction (we have assumed that peak and valley voltages are the same in all semiconductors, which is, of course, an approximation); p can thus vary rapidly with N and P. Approximately we have

$$p = \exp\left\{ -\alpha m_r^{1/2} E_g \left(\frac{1}{N} + \frac{1}{P} \right)^{1/2} \right\}$$

where α is a constant and m_r, the reduced mass, is defined by

$$\frac{1}{m_r} = \frac{1}{m_e} + \frac{1}{m_{hl}}.$$

Thus, if N and P can be made large enough in a given semiconductor, a fast tunnel diode can be made with that semiconductor. It will be noticed that if E_g and m_r are small, N and P do not need to be so great for comparable speed as they do when E_g and m_r are large. A further point of importance in fabrication is that the solubilities must be large enough at a temperature for which the impurity diffusion coefficient is low. Tunnel diodes are usually made by the alloying process; at the elevated temperature involved the impurity atoms can diffuse or drift in the junction field, if one still exists at the fabrication temperature; this will increase the junction width over the minimum attainable for a given N and P (which will occur when there is a sharp 'step' junction); the tunnel current will then be reduced and the device will be slower electronically. Since impurity diffusion coefficients in solids increase sharply as the temperature increases these considerations imply a limitation on the temperature cycle used during alloying; a long time at too high a temperature will degrade the junction.

From what has been said above it is clear that the attainable solubilities of donor and acceptor impurities are of great importance, as are their diffusion coefficients. It was for this reason that InSb, for which the solubility requirements

* Detailed accounts are available elsewhere (see Sommers 1959, and accompanying article by Jonscher).

for a very fast electronic speed are satisfied very easily, has attracted interest. When Batdorf *et al.* (1960) first described very high speed InSb tunnel diodes, there was some doubt as to whether other semiconductors such as Ge and Si could be used to make such extremely fast diodes (the value of C/G can easily be made as small as 5×10^{-12} s). The present situation appears to be that Ge (and GaAs) diodes can be made fast enough for most practical requirements, whilst Si diodes seem to remain less promising. Ge and GaAs diodes do not require cooling (although they will work quite satisfactorily if cooled), so that InSb diodes are at an immediate disadvantage, unless cooling were to be used in the circuitry in another connection, e.g. for a superconducting computer store. In applications as a negative resistance amplifier it should be possible to obtain lower noise figures using InSb than would be obtained with tunnel diodes of the other semiconductors (Chang 1960). The basic reason for this is that the voltages at which peak and valley occur are much lower for InSb (typically 25 mv and 125 mv, respectively) than for Ge (75 mv and 350 mv) or GaAs (120 mv and 600 mv). However, it seems doubtful whether the noise figure improvement would justify the need for cooling in many applications. In summary, it now seems less likely that InSb tunnel diodes will be widely used in circuitry.

There is, however, much interest in InSb tunnel diodes as a tool for studying the tunnelling process. The reason for this is that they are more susceptible to theoretical analysis than tunnel diodes made of Ge, because the band edges are at the centre of the Brillouin zone, and phonons are not required in tunnelling transitions. Also, several experiments are possible on InSb tunnel diodes which are difficult on Ge and GaAs tunnel diodes. For example, the tunnel current can be reduced appreciably by large magnetic fields. Calawa *et al.* (1960) have reported measurements of the variation of tunnelling current with applied magnetic field at 77° K, and the phenomenon has been studied in more detail by Butcher *et al.* (1961). Fig. 1 shows the variation of peak current density J_p with magnetic field H when the tunnelling current is parallel to the magnetic field. The full line represents the theoretical variation

$$\frac{J_p(H)}{J_p(O)} = \frac{H \cosh (gm_r H/m_0 H_0)}{H_0 \sinh (H/H_0)} .$$

The experimental results have been fitted to this curve by finding the value of H_0 for each diode needed to give the best fit. In so doing the values $g = 56$ and $m_r/m_0 = 0{\cdot}007$ were used, as is appropriate for InSb. By developing the theory of tunnelling through a parabolic potential barrier (as opposed to the linear potential barrier usually assumed), theoretical values of H_0 were derived by Butcher *et al.* (1961) without the use of any adjustable parameters. It was then possible to compare the experimental and theoretical values of H_0, as is done in the Table. It will be seen that theory

Magnetic field scaling factors for peak currents

Diode	*Experimental* H_0 *(kOe)*	*Theoretical* H_0 *(kOe)*
A	16	20
B	19	29
C	23	35

and experiment are in reasonable agreement, despite the rather drastic simplifications which must still be made if the theory is to remain mathematically tractable. Another test of the parabolic barrier theory is possible. In Fig. 2 are shown data obtained on the peak current of diodes made on material with various donor concentrations; each point is the averaged result for a number of diodes made from

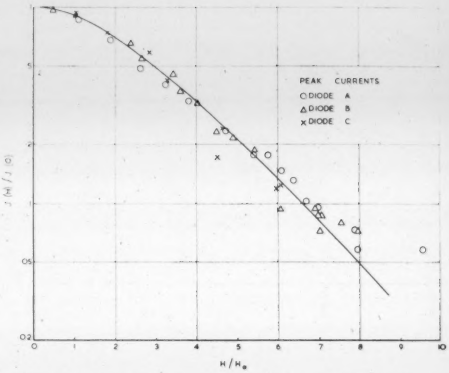


Fig. 1. Ratio of peak current density with and without applied magnetic field plotted against the reduced magnetic field for three InSb tunnel diodes with different values of the doping concentration N. The full line is the theoretical variation given by the equation in the text, and the experimental points have been fitted to the curve by adjusting the values of H_0. The experimental values of H_0 thus found are compared with the theoretical values in the table. (*From* Butcher *et al.* 1961.)

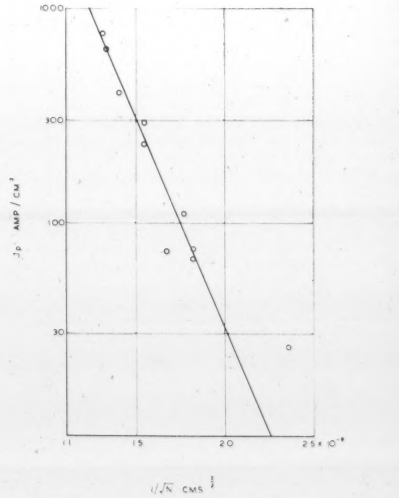


Fig. 2. Variation of peak current density with the doping concentration N. Each point gives the average result for a number of diodes made on material with the same value of N. The line has been drawn through the experimental points bearing in mind their approximate statistical weight. (*From* Butcher *et al.* 1961.)

material with the same donor concentration. From the expression given above for p, we expect a plot of $\log J_p(0)$ against $(1/\sqrt{N})$ to be almost linear provided $P \gg N$. The line in the figure has been drawn as the best fit to the experimental points, bearing in mind their statistical weight. The detailed theory actually requires that the points should be plotted on more complicated axes; the theoretical parameters describing the slope and position of the line are then $3{\cdot}7 \times 10^9\ \mathrm{cm}^{-3/2}$ and $3{\cdot}9 \times 10^{-4}\ \mathrm{A\,cm}^{-1/2}$ respectively. The experimental parameters on the altered plot are $4{\cdot}3 \times 10^9\ \mathrm{cm}^{-3/2}$ and $0{\cdot}60\ \mathrm{A\,cm}^{-1/2}$. Here again the theory is in reasonable agreement with experiment. For all these experiments the diodes were made by alloying small disks of pure Cd on to Te doped InSb, using a peak alloying temperature of about 400° C; the Te doped crystals were grown on a ⟨110⟩ growth axis to avoid the non-uniformities associated with the 'facet effect' (see Hulme and Mullin 1959 and Mullin and Hulme 1960). Te is a donor impurity in InSb and Cd is an acceptor. The fabrication conditions made it reasonable to assume $P \gg N$.

The agreement obtained between theory and experiment is taken to indicate that the theoretical concepts used in explaining the tunnel diode characteristics are basically sound. In the theory it was assumed that tunnelling involved only the light-hole band; this assumption has been questioned by Chynoweth *et al.* (1960), but Butcher *et al.* (1961) would support Haering and Miller (1961) in upholding the assumption. It might be mentioned that the effect of magnetic fields on Ge diodes is very small (a few per cent at 150 kOe). Even for InSb, where the material parameters are favourable, the effect is not, unfortunately, large enough to form the basis of a magnetically controlled tunnel diode. To obtain a tunnel diode for which magnetic fields of a more conventional size would provide an effective 'third terminal', one would have to turn to a semiconductor with an even smaller forbidden energy gap; operation at 77° K or below would be an almost inevitable concomitant.

Close studies of the voltage–current curves (or better, the slope conductance as a function of voltage) at 4° K have revealed detail of fundamental significance. A current contribution associated with phonon-assisted transitions has been observed by Hall *et al.* (1960) and others. We have already remarked that the band edges in InSb are at the centre of the Brillouin zone; there is thus no large change in the wave vector (i.e. value of **k**) of an electron involved in a tunnelling transition. Transitions can therefore occur without the need for the participation of a phonon to conserve wave vectors in the transition. A process involving a phonon with *zero* wave vector is not excluded by this reasoning, although it will be rather less likely than the process not involving a phonon. Acoustical phonons with zero wave vector have zero energy, but optical phonons of zero wave vector have a finite energy. Now, at 4·2° K there are virtually no phonons present, and processes involving phonon absorption are excluded, but the process involving optical phonon emission will be possible when the bias voltage gets high enough to provide the necessary energy. In fact, the onset of the tunnelling process involving phonon emission has been observed by Hall *et al.* (1960); it is marked by a change in the slope of the graph of dI/dV against V. The threshold voltage is 24 mv which is in good agreement with the optical phonon energy deduced from other experiments.

Hall *et al.* (1960) also observed detail in the dI/dV against V curve near zero bias. This was attributed to polaron effects and allowed the polaron formation energy to be estimated. (A polaron consists of an electron associated with the lattice deformation it causes by electrostatically attracting and repelling lattice atoms with positive and negative charges respectively; polarons are not possible in elemental semiconductors.)

It is thus clear that the study of InSb tunnel diodes has already contributed to a more detailed understanding of the physical processes operative in tunnelling.

Acknowledgments

I would like to acknowledge my indebtedness to colleagues at the Royal Radar Establishment for stimulating discussions on various aspects of tunnel diodes.

References

BATDORF, R. L., DACEY, G. C., WALLACE, R. L., and WALSH, D. J., 1960, *J. Appl. Phys.*, **31**, 613.

BUTCHER, P. N., HULBERT, J. A., and HULME, K. F., 1961, *J. Phys. Chem. Solids*, in the press.

CALAWA, A. R., REDIKER, R. H., LAX, B., and MCWHORTER, A. L., 1960, *Phys. Rev. Letters*, **5**, 55.

CHANG, K. K. N., 1960, *Proc. Inst. Radio Engrs, N.Y.*, **48**, 107.

CHYNOWETH, A. G., LOGAN, R. A., and WOLFF, P. A., 1960, *Phys. Rev. Letters*, **5**, 548.

HAERING, R. R., and MILLER, P. B., 1961, *Phys. Rev. Letters*, **6**, 269.

HALL, R. N., RACETTE, J. H., and EHRENREICH, H., 1960, *Phys. Rev. Letters*, **4**, 456.

HULME, K. F., and MULLIN, J. B., 1959, *Phil. Mag.*, **4**, 1286.

MULLIN, J. B., and HULME, K. F., 1960, *J. Phys. Chem. Solids*, **17**, 1.

SOMMERS, H. S., 1959, *Proc. Inst. Radio Engrs, N.Y.*, **47**, 1201.

WELKER, H., 1952, *Z. f. Naturf.*, **7a**, 744.

The physics of the tunnel diode

by A. K. JONSCHER, B.Sc., Ph.D., A.M.I.E.E., A.Inst.P., The General Electric Company Limited, Central Research Laboratories, Hirst Research Centre, Wembley, Middlesex

MS. received 12th April 1961; paper presented at the Symposium on Tunnel Diodes held on 7th February 1961

Abstract

A brief account is given of the physical principles governing the operation of tunnel diodes. The material parameters affecting the peak and valley currents and the capacitance are discussed. The significance of some low temperature phenomena is explained.

THE invention of the transistor in 1948 stimulated intensive scientific and technological activity on semiconducting materials and on devices based on the principle of injection of excess carriers by p–n junctions. The main emphasis was on high-purity semiconductors as starting material for improved devices. This trend naturally favoured silicon and germanium and during the following decade the understanding of these materials reached a very advanced stage.

In consequence, the focus of scientific interest began to shift perceptibly to other materials and to different phenomena, further refinements of diodes and transistors being considered to lie in the sphere of technology rather than of science.

Just at that time, in 1957, the tunnel diode was invented by the Japanese physicist Leo Esaki in the course of his work on breakdown in very heavily doped materials. This device was quickly recognized as having great possibilities and considerable effort was switched to its development in laboratories throughout the world. The stage was set for very rapid development of ideas and techniques and today the output of published information on tunnel diodes is correspondingly high.

The interest in the tunnel diode as a very high frequency negative resistance device and a very fast switch hardly needs stressing in the present context. What may be less clearly appreciated are the great opportunities for solid state research opened up by the emergence of the tunnel diode. The device employs semiconductor materials which are doped with foreign impurities to the limit to which the crystal lattice will accept them without disintegrating. This poses severe problems in the growing and alloying operations. Secondly, the operation of the device is uniquely sensitive to the nature of the band structure of the material employed and the study of device characteristics affords a means of understanding the structure of heavily doped materials. This is precisely the aspect, however, about which our present knowledge is singularly inadequate—very little is known about the manner in which the incorporation of high densities of impurities modifies the intrinsic energy bands.

The operation of the tunnel diode is based on the quantum mechanical tunnel effect—a well-known phenomenon invoked more than a quarter of a century ago to explain field emission from solids and breakdown in dielectrics. With reference to Fig. 1, consider an electron of energy ϵ and charge $-q$ confronted with a potential barrier $qV(x)$ of arbitrary shape such that $qV > \epsilon$ over part of the barrier width. From the standpoint of classical mechanics the barrier is impervious to the electron which becomes totally reflected on it. Quantum mechanically, however, the electron

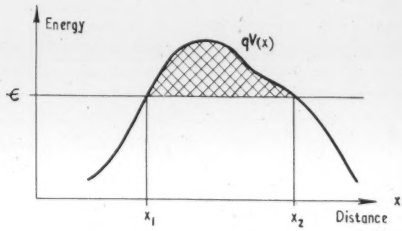


Fig. 1. Quantum mechanical tunnelling through a potential barrier.

has a finite probability of penetrating a barrier higher than its own energy. This probability is given by the expression

$$W = \exp\left\{-2\int_{x_1}^{x_2}\frac{(2m)^{1/2}}{\hbar}(qV-\epsilon)^{1/2}dx\right\} \qquad (1)$$

where $x_1 - x_2$ is the interval in space over which the energy of the barrier exceeds that of the electron, m is the mass of the electron and $\hbar = 2\pi h$ is Planck's constant. The strong dependence of the probability of transmission on the width and the height of the barrier is immediately apparent.

The situation arising in the case of a solid insulator to which a high electric field E is applied is shown in Fig. 2. The energy bands are sloping and a valence band electron at a level ϵ faces a triangular energy barrier of height ϵ_G and thickness $\epsilon_G/(qE)$, where ϵ_G is the width of the forbidden gap. On the other side of the barrier are the allowed empty levels of the conduction band. The relevant expression for the probability of tunnelling is

$$W = \exp\left\{-\frac{4}{3}\frac{(2m^*)^{1/2}}{\hbar q}\frac{\epsilon_G^{3/2}}{E}\right\} \qquad (2)$$

where m^* is an appropriate effective mass.

p–n junctions in semiconductors contain regions in which a high internal electric field exists even in the absence of an external bias. Fig. 3(*a*) shows the equilibrium energy band diagram of a p–n junction in which both regions are non-degenerate, i.e. the Fermi level ϵ_F falls within the forbidden gap at a distance in energy of ζ_n and ζ_p from the edges of the

conduction and valence bands respectively. In this case the equilibrium carrier densities n_0 and p_0 in the n- and p-regions are given by the expressions

$$n_0 = N_c \exp(-\zeta_n/kT),\ p_0 = N_v \exp(-\zeta_p/kT) \qquad (3)$$

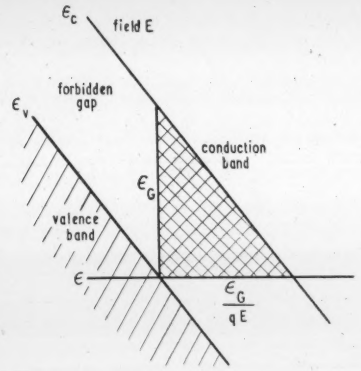


Fig. 2. Interband transition in a solid by tunnelling in a high electric field.

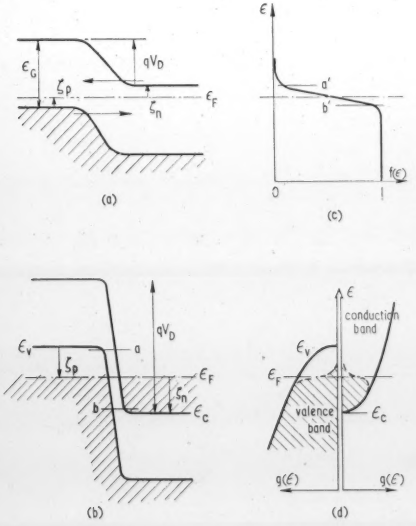


Fig. 3. Tunnelling transitions in p–n junctions in equilibrium.
(*a*) Non-degenerate p–n junction—no transitions are possible.
(*b*) p–n junction with degenerate regions on either side—finite transition probabilities exist in the range of energies between a and b.
(*c*) Fermi–Dirac distribution function showing the range a′b′ in which the product $f(1-f)$ has an appreciable magnitude.
(*d*) Density of states and electron density distribution in energy.

where N_c and N_v are the effective densities of states defined as

$$N_{c,v} = 2\left(\frac{2\pi m^*_{e,h} kT}{h^2}\right)^{3/2}.$$

m_e^* and m_h^* are the appropriate density-of-states effective masses for electrons and holes, respectively, k is Boltzmann's constant and T the absolute temperature. For silicon and germanium at room temperature N_c and N_v are approximately equal to 10^{19} cm^{-3}. Correspondingly lower values apply at lower temperatures and in materials with low effective masses, such as *n*-type InSb. A semiconductor is non-degenerate if the donor impurity density N_D or the acceptor impurity density N_A is sufficiently low for the resulting free carrier density n_0 or p_0 to be substantially less than N_c or N_v respectively.

The introduction of a higher impurity density causes the Fermi level to fall effectively within the conduction or the valence band, as shown in Fig. 3(*b*), where the p- and n-regions on either side of the junction are degenerate. The density of electrons at an energy level ϵ is given by the product $g(\epsilon)f(\epsilon)$, where $g(\epsilon)$ is a density-of-states function and $f(\epsilon)$ is the Fermi–Dirac distribution function

$$f(\epsilon) = [1 + \exp\{(\epsilon - \epsilon_F)/kT\}]^{-1}. \qquad (4)$$

The function $f(\epsilon)$ is shown diagrammatically in Fig. 3(*c*) and the density-of-states function $g(\epsilon)$ for the simplest possible case is shown in Fig. 3(*d*). The latter diagram also shows as the dotted line the product $g(\epsilon)f(\epsilon)$. The shaded areas correspond to levels occupied by electrons, while the open space near the top of the valence band refers to the hole population.

No tunnelling transitions are possible in a non-degenerate p–n junction in equilibrium or at a slight forward or reverse bias. The reason for this is that the tunnelling process normally conserves energy and therefore corresponds to horizontal transitions in the energy band diagram. No electron from the valence band in Fig. 3(*a*) can make a tunnelling transition into the conduction band along the horizontal arrow. Similarly, no electron can tunnel from the conduction band on the n side into the valence band. The latter process, in any case, would be very unlikely in a p–n junction in which at least one side is non-degenerate, even if the energy difference could somehow be absorbed. The reason for this is that the actual rate of transitions contains the product of a transition probability, the number of electrons available for transition and the number of empty states (holes) which can receive these electrons. The product of electron density and hole density is, however, very small in such a p–n junction. Tunnelling becomes possible on the application of a reverse bias (Fig. 4) when the maximum field is sufficiently high to give an appreciable probability of tunnelling according to Eqn (2). This mechanism is responsible for breakdown in p–n junctions that are fairly heavily doped on both sides and are yet non-degenerate. The magnitude of the maximum electric field in an abrupt p–n junction is given by the expression

$$E_{max} = 2\left(\frac{2\pi q}{\kappa}\right)^{1/2}(V_D - V)^{1/2}\left(\frac{N_D N_A}{N_D + N_A}\right)^{1/2} \qquad (5)$$

where κ is the dielectric constant, V is the applied bias reckoned positive in the forward direction and V_D is the height of the equilibrium potential barrier as shown in Fig. 3.

In a p–n junction in which both regions are degenerate, as shown in Fig. 3(*b*), tunnelling transitions are possible in

equilibrium in the energy range between a and b. Transitions will actually occur only over a smaller interval in which both electrons and holes are available in reasonable numbers, that is in the range where the product $f(\epsilon)[1-f(\epsilon)]$ is sufficiently large. This range is indicated by a′b′ in Fig. 3(*c*). There is no net flow of current in equilibrium since transitions occur with equal frequency in both directions.

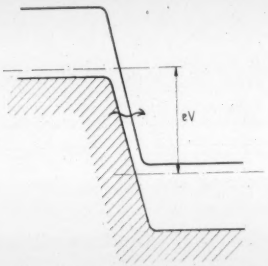


Fig. 4. Tunnelling in a reverse biased non-degenerate p–n junction.

The situation arising upon the application of a small forward bias is shown in Fig. 5(*a*). The conduction band is displaced upwards in energy with respect to the valence band. This enhances the electronic transitions from right to left and inhibits the opposite flow, the result being a sharp rise of current sketched on the V–J characteristic on the right. With increasing bias, however, the overlap between filled levels in the conduction band and empty levels in the valence band decreases and this results in the fall of current, Fig. 5(*b*). Eventually a bias is reached at which the overlap between bands is reduced to zero (Fig. 5(*c*)) and the direct tunnelling current should fall to zero since no transition at constant energy can take an electron from one side to the other. The peak current J_p and corresponding voltage V_p form one set of important device parameters.

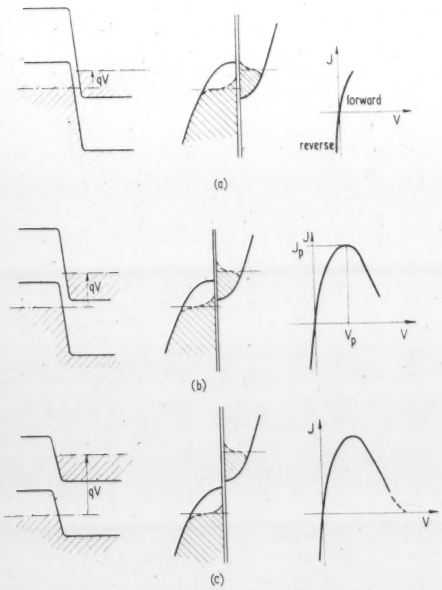


Fig. 5. Energy band diagram, electron density distributions and voltage current characteristic of a degenerate p–n junction at a small forward bias.

The peak voltage V_p is determined by the position of the Fermi level relative to the band edges and hence depends on the impurity density. It is virtually independent of the magnitude of the forbidden gap. It is not possible to specify quantitatively the value of the peak current J_p, but the following expression for the zero-bias conductance $(dJ/dV)_0$ gives the dependence of the current density on the principal physical parameters:

$$\left(\frac{dJ}{dV}\right)_0 = 0{\cdot}16\left(\frac{Nm_r/m_0}{\kappa}\right)^{1/2}\exp\left\{-5{\cdot}8\times10^{10}\left(\frac{\epsilon_G{}^2\kappa m_r/m_0}{N}\right)^{1/2}\right\}. \tag{6}$$

Here J is in amperes/cm^2, V in volts, ϵ_G in electron volts. N is the impurity density in the less heavily doped region (assuming that the other region has a substantially higher impurity density). m_r is a reduced effective mass given by

$$\frac{1}{m_r} = \frac{1}{m_e{}^*} + \frac{1}{m_h{}^*} \tag{7}$$

and $m_h{}^*$ refers to the light holes where appropriate. m_0 is the mass of a free electron. The very critical dependence of $(dJ/dV)_0$ upon ϵ_G, m_r/m_0 and N can be seen from the graphs in Fig. 6 which refer to three sets of parameters corresponding to InSb, Ge and Si.

The practical importance of operating at a high peak current density arises from the fact that this helps to reduce the area of the junction and therefore the capacitance. The

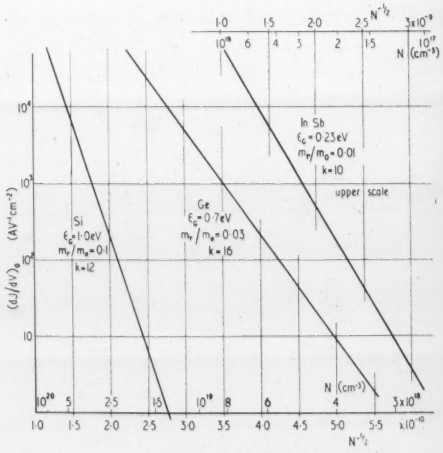


Fig. 6. Computed zero bias conductance $(dJ/dV)_0$ as a function of the impurity density for Si, Ge and InSb.

magnitude of the latter per unit area of the junction is given by the expression

$$C = 0{\cdot}80 \times 10^{-4}\left(\frac{\kappa N}{\epsilon_G}\right)^{1/2} \text{pF/cm}^2. \qquad (8)$$

In view of the very high values of N the capacitance is high by usual semiconductor device standards.

The small-signal equivalent circuit of a tunnel diode is shown in Fig. 7. It consists of a negative conductance $-G = -1/R$ in parallel with the junction capacitance C.

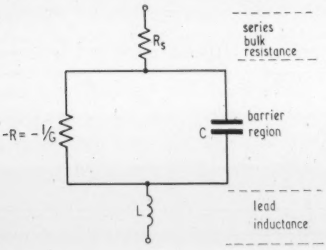


Fig. 7. Equivalent circuit of a tunnel diode.

The series resistance R_s is due to the bulk material of the device on either side of the junction. The lead inductance L can represent an important limitation in view of the high operating frequencies.

The complete voltage–current characteristic of a tunnel diode is shown in Fig. 8. Two points should be noted.

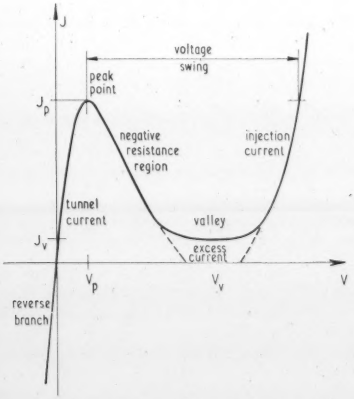


Fig. 8. Complete voltage current characteristic of a tunnel diode.

Firstly, the current does not go down to zero as expected from simple theory, but remains at a finite value in the valley region of the characteristic. At still higher voltages, comparable with the energy gap, the characteristic shows a steep rise due to the ordinary injection process in the junction. The valley current J_v is not due to this mechanism and is therefore referred to as the excess current.

The ratio of peak current to valley current is reported to be sensibly independent of temperature, pressure and impurity density. On the other hand, this ratio varies appreciably with the material and with the nature of the impurity. The cause of the excess current appears to have been established recently as tunnelling transitions via some energy levels in the forbidden gap (Fig. 9). The origin of these levels is at

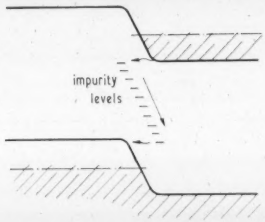


Fig. 9. Tunnelling through intermediate localized levels in the valley region of the characteristic.

present unexplained—they are thought to be due to 'band edge tails' inherent in heavily doped materials. Insufficient is known of the impurity levels in heavily doped materials to suggest even broad models for this mechanism. It is interesting to know, however, that the ratio J_p/J_v may be as high as 60 in gallium arsenide, while a typical figure for germanium is 10 and for silicon 5 or less.

Further refinement of the tunnelling mechanism is called for in the case of materials having complicated band structures. The simple central parabolic bands, shown in Fig. 10(*a*),

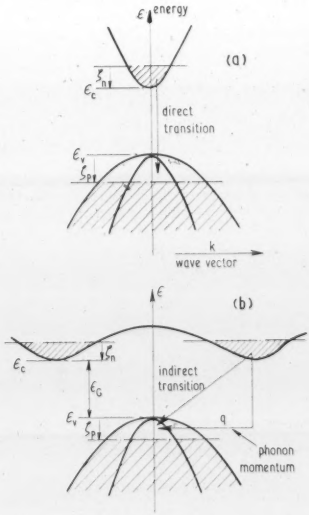


Fig. 10. Direct transitions and indirect phonon assisted transitions in typical semiconductor band structures.

have the minima and maxima of energy at zero wave vector k in the reduced Brillouin zone. By raising the valence band with respect to the conduction band it is possible to get direct tunnelling transitions under simultaneous conservation of both energy and momentum. This type of band structure is observed in some III–V compounds, such as GaAs. On the other hand, germanium, silicon and probably some III–V compounds such as GaP exhibit a more complicated band structure in which the energy minima of the conduction band fall at non-zero k values, as shown in Fig. 10(*b*). The conservation of momentum requires in this instance the emission of a lattice vibration quantum (phonon) to dissipate the momentum difference q involved in a tunnelling transition. This in turn slightly upsets the energy balance since the phonon has a definite energy of its own. The result is that a graph of the conductance dJ/dV against V for a tunnel diode at very low temperatures shows a definite structure in which the peaks can be identified with phonon energies (Fig. 11).

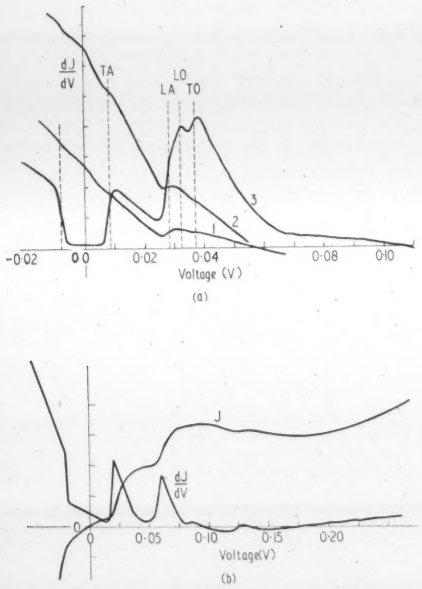


Fig. 11. The voltage dependence of current J and conductance dJ/dV at a low forward bias in tunnel junctions at 4·2° K showing structure due to phonon interactions.

(*a*) Germanium junctions formed by alloying 1% Ga, 99% In to n-type germanium of the following impurity densities:

Curve	*Donor*	*Concentration*
1	P	13×10^{18} cm^{-3}
2	As	9×10^{18} cm^{-3}
3	Sb	14×10^{18} cm^{-3}

The vertical lines denote phonon energies as follows: TA, transverse acoustical; LA, longitudinal acoustical; TO, transverse optical; LO, longitudinal optical.

(*b*) Silicon junction formed by alloying boron doped aluminium to silicon containing 2×10^{19} cm^{-3} of phosphorus.

(From R. N. Hall, *Proceedings of the International Conference on Physics of Semiconductors, Prague*, 1960.)

A detailed study of this structure in devices using different donor impurities shows that the nature of the impurity has a profound effect on the phonon interactions and thus betrays modifications of the band structure by the doping impurity. Thus, it appears that Sb-doped germanium shows very pronounced phonon structure, but P- and As-doped germanium shows very little. This is interpreted as an indication of the predominance of direct transitions due to a substantial lowering of a subsidiary minimum of the conduction band at $k = 0$ under the influence of heavy doping.

Not unexpectedly the tunnelling process is affected by the presence of magnetic fields in the range 10^4–10^5 gauss, both parallel and normal to the junction plane. The effects of the field on the motion of the carriers and on the band structure combine to give an overall lowering of the current throughout the whole range of the characteristic. With the magnetic field H parallel to current flow the following expression for the ratio of current with and without the field appears to be satisfied

$$\log\left\{\frac{I(H)}{HI(0)}\right\} \propto -H.$$

Some oscillatory phenomena have also been reported and attributed to the Landau splitting of energy levels. All magneto-tunnelling effects are most pronounced in indium antimonide due to the low effective mass of carriers in this material.

It is clear that further detailed study of this subject will throw badly needed light on the very complex problem of interaction between the impurities and the host lattice in semiconductors.

The development of tunnel diodes in various materials has also stimulated interest in the growing of extremely heavily doped single crystal materials. Our present understanding of this important subject is rather rudimentary and advances in this direction are bound to be of interest from both the applied and the theoretical standpoint.

Bibliography

Original publication

ESAKI, L., 1958, *Phys. Rev.*, **109**, 603.

General and review papers

ESAKI, L., and MIYAHARA, Y., 1960, *Solid State Electronics*, **1**, 13.

HALL, R. N., 1960, *I.R.E. Trans on Electron Devices*, p. 1 (January).

LESK, I. A., HOLONYAK, N., DAVIDSOHN, A. S., AARONS, M. W., 1959, *I.R.E. Wescon Convention Record*, Part 3, p. 9.

LESK, I. A., and SURAN, J. J., 1960, *Electrical Engineering* (April).

PUCEL, R. A., 1960, *Solid State Electronics*, **1**, 22.

SOMERS, H. S., Jr., 1959, *Proc. I.R.E.*, **47**, 1201.

Phonon-assisted tunnelling, excess current and other physical phenomena

CHYNOWETH, A. G., FELDMAN, W. A., and LOGAN, R. A., 1961, *Phys. Rev.*, **121**, 684.

CHYNOWETH, A. G., WANNIER, G. H., LOGAN, R. A., and THOMAS, D. E., 1960, *Phys. Rev. Letters*, **5**, 57.

HALL, R. N., 1960, *Proceedings of the International Conference on Physics of Semiconductors, Prague* p. 193.

HOLONYAK, N., LESK, A. I., HALL, R. N., TIEMANN, J. J., and EHRENREICH, H., 1959, *Phys. Rev. Letters*, **3**, 167.

KANE, E. O., 1959, *J. Phys. Chem. Solids*, **12**, 181.
—— 1961, *J. Appl. Phys.*, **32**, 83.
MORGAN, J. V., and KANE, E. O., 1959, *Phys. Rev. Letters*, **3**, 466.

Tunnel diodes in III–V compound semiconductors

BATDORF, R. L., DACEY, G. C., WALLACE, R. L., and WALSH, D. J., 1960, *J. Appl. Phys.*, **31**, 613.
BURRUS, C. A., 1961, *J. Appl. Phys.*, **32**, 1031.
CALAWA, A. R., REDIKER, R. H., LAX, B., and MCWHORTER, A. L., 1960, *Phys. Rev. Letters*, **5**, 55.
CHYNOWETH, A. G., LOGAN, R. A., and WOLFF, P. A., 1960, *Phys. Rev. Letters*, **5**, 548.
HALL, R. N. RACETTE, J. H., and EHRENREICH, H., 1960, *Phys. Rev. Letters*, **4**, 456.
HOLONYAK, N., and LESK, I. A., 1960, *Proc. I.R.E.*, **48**, 1405.
KLEINKNECHT, H. P., 1961, *Solid State Electronics*, **2**, 133.

Magneto-tunnelling phenomena

CALAWA, A. R., REDIKER, R. H., LAX, B., and MCWHORTER, 1960, *Phys. Rev. Letters*, **5**, 55.
CHYNOWETH, A. G., LOGAN, R. A., and WOLFF, P. A., 1960, *Phys. Rev. Letters*, **5**, 548.
HAERING, R. R., and ADAMS, E. N., 1961, *J. Phys. Chem. Solids*, **19**, 8.

Electroluminescent devices

by H. K. HENISCH, Ph.D., F.Inst.P., Department of Physics, The University, Reading, Berks.

MS. received 9th June 1961

Abstract

A survey is made of recent progress in the applications of electroluminescence, dealing with panels for illumination purposes, with alphanumerical indicators, with light and picture amplifiers of various kinds, as well as scanned picture displays.

1. Introduction

THE development of electroluminescence began in 1923 with the discovery of light emission (Lossew 1923) from silicon carbide detector contacts. At the time this was an intriguing but only an isolated phenomenon. In its various manifestations, it has since come to be regarded as a common property of non-metallic solids and has given rise to a whole series of devices of great potential importance. The simplest of these are now commercially available and various others have been made in the form of demonstration models. Others still can be found only in review articles on electroluminescence (!) but, as far as the future is concerned, these may well be among the most important. The purpose of this paper is to give a brief account of current work and to survey the range of interesting possibilities which this field has to offer. The subject has an extensive patent literature to which a bibliography recently prepared by Ivey (1959, 1961) is a valuable guide.

2. Electroluminescent panels for illumination purposes

Electroluminescent panels are made in several forms, of which three are illustrated schematically in Fig. 1. In each case an alternating voltage of power or audio frequency is applied between the two conducting layers. The phosphor which is suspended in an organic dielectric (Figs 1(*a*) and (*b*)) or in a ceramic medium (Fig. 1(*c*)), is in a high field which causes it to emit light. In all cases, the surface brightness increases sharply with increasing applied voltage, in the general manner shown by curve A of Fig. 3. Empirical expressions are available which describe these relationships with astonishing success (Zalm *et al.* 1955), considering the inherent complications of such a powder system. A great deal of research on the mechanism of this process has been done during the last few years and, accordingly, a framework of theoretical interpretation is now available. In most of the materials which are here relevant electroluminescence involves, in all likelihood, the impact ionization of activator centres by fast electrons. On the other hand, much of our knowledge is still empirical and many important questions of detail remain unanswered.*

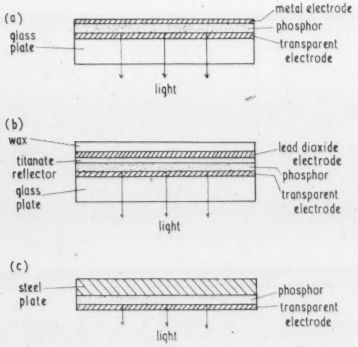


Fig. 1. Electroluminescent light sources (*a*) *after* Mager (1951), (*b*) *after* Mash (personal communication), (*c*) *after* Rulon (1955).

The apparently simple structures on Fig. 1 contain a large number of variable parameters which have to be optimized to achieve high brightness, long working life, the desired colour response and satisfactory power efficiency. As one would expect, these aims lead to conflicting requirements, and compromises have to be made to suit specific applications. Glass panels which make use of a phosphor embedded in an organic plastic are, on the whole, the most efficient, whereas panels which are based on steel substrates obviously have greater mechanical strength. It is possible, for instance, to puncture such a steel plate with a rifle bullet without impairing the emission of the surrounding area, an unorthodox but convincing demonstration of robustness. The ceramic matrix offers improved protection against atmospheric influences which otherwise cause the phosphor to deteriorate. It does, however, lead to a somewhat lower light output. The best power efficiencies now quoted (15 lm/w, for organic dielectrics) equal the values which are typical for filament lamps of about 100 w consumption. The dielectric matrix inevitably absorbs some of the light and we may therefore conclude that the intrinsic efficiency of the phosphor itself is higher. The quoted values have been increasing over a number of years and there is no reason for believing that an ultimate limit has been reached. During operation the panels become only slightly heated, by virtue of their large dissipating surface. The panel on Fig. 1(*b*) is a development of the version in Fig. 1(*a*). The white titanate reflector serves to increase the effective light output and a lead dioxide layer is used as one of the electrodes because of its 'self-sealing' properties. It is a semiconductor which changes composition on heating and thereby becomes less conducting. A weak spot in the plate arising from faulty manufacture would initially draw a high current, but would do so for a short

* For a general treatment, see *Electroluminescence* by H. K. Henisch; International Series of Monographs on Semiconductors (London: Pergamon Press), to be published.

period only until the situation is corrected by the heat developed and by the corresponding (permanent) increase of electrode resistance. The wax layer is for protection against moisture.

As far as life-expectancy is concerned, these panels differ greatly from filament lamps. When a filament breaks, the consequences are unambiguous. In contrast, an electroluminescent panel rarely 'goes out', it merely deteriorates, giving a diminishing light output during a prolonged period of use. An estimate of its life is therefore very much dependent on one's definition of 'death'. Instances are known in which the useful life, based on reasonable criteria, has been of the order of 50 000 hours, but the units which are now commercially available deteriorate more quickly. These matters are again governed by a compromise. By applying higher voltages, it is possible to achieve a higher surface brightness, but always at the cost of lowering the life expectancy. For some applications, and specifically for aircraft, the absence of catastrophic failures could be a useful safety advantage over conventional devices.

For ordinary illumination purposes one of the main advantages of electroluminescent panels is their virtually two-dimensional character. The principal disadvantage is their low surface brightness. Although surface brightness values up to 1000 ft-lamberts have been claimed for very short periods on experimental units, typical operational levels of commercially available panels are very much lower. Common values are between 1 and 10 ft-lamberts. There are, however, quite a number of applications (architectural tiles, darkroom lamps, signboards, etc.) for which high brightness is not required. A potentially important and large-scale use of electroluminescent panels is in the field of road signs with self-luminous lettering.

In the normal way, using phosphors based on zinc sulphide, the emission is in the visible region of the spectrum. Larach and Shrader (1956) have shown that ultra-violet emission can be achieved by the use of boron nitride. The efficiencies obtainable are not yet known. If they were at all reasonable the process could lead to the design of cool area sources of ultra-violet radiation, and such devices could be used for blue-printing and related processes.

The panels which are now commonly made are either rigid (if based on glass) or only slightly flexible (if based on steel). In addition electroluminescent light sources have been described which are intended to be highly flexible (Gillson 1956). These can be made by dispersing the phosphor in a plastic film and applying semi-transparent conducting electrodes on each side, but the practical difficulties of manufacture are considerable.

3. Alphanumerical and block pattern displays

It is evidently possible to construct electroluminescent panels with multiple electrodes which can be selectively energized. Three such examples are shown in Fig. 2. In the case of glass-based panels it is the metal back electrode which is shaped in some such way; in the case of metal-based panels it is the transparent top electrode. By energizing some of the areas in combination, recognizable letters and numerals can be formed. The larger the number of electrodes, the more satisfactory is the shape of the symbols, though at the cost of complicating the switching system. The selective operation of electrodes in any case constitutes somewhat of a problem. From the circuit point of view, a panel designed for numerical information only (Fig. 2(*b*)) should be capable of being operated by a single-pole ten-way switch. This cannot be done by direct connections from switch to panel, but necessitates an intermediate stage. For this purpose, Mash (1960) has designed a 'coding matrix' in the form of a flat plate which can be co-planar with the panel or can be mounted elsewhere as a separate component. The external switch is connected to the coding matrix and is used to select one of ten different circuit patterns. Since no two electrodes can be allowed to remain permanently connected, the coding matrix must contain isolating components which are voltage dependent. Mash used silicon carbide elements which appear in series connection with the panel electrodes. Their function is to increase the non-linearity of the brightness–voltage relationship. Fig. 3 illustrates this. Curve A shows

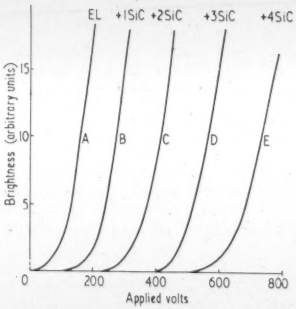


Fig. 3. Brightness–voltage relationships for an electroluminescent element and non-linear series components (*after* Mash 1960).

a relationship which is typical of an electroluminescent panel alone, curves B–E refer to a panel with one or more silicon carbide elements in series. The much sharper onset of brightness arises, of course, from the fact that the series resistance is very high at low voltages and collapses quickly as the voltage increases. The outcome is the appearance of what is almost a critical threshold voltage for luminescence. This greatly simplifies the interconnections for switching, not only for alphanumerical displays but also in other contexts. In principle, junction diodes with sharply defined breakdown characteristics could achieve this even more effectively, but a bias voltage would normally be necessary and this would destroy some of the simplicity of the series arrangement.

Panels of the type shown on Fig. 2(*c*) can be used as alphanumerical displays, but are unduly complex for this purpose alone. They are suitable for the display of simple pictorial information and of coded block diagrams. Another form of block display involves the use of two crossed arrays of strip electrodes, as shown on Fig. 4. When a sufficiently high voltage is applied between two strips, the phosphor under the intersection will emit light, and this light spot can

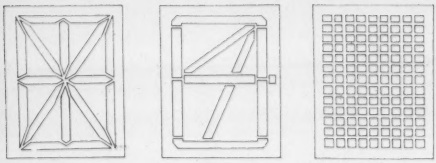


Fig. 2. Multi-electrode information display panels.

be moved arbitrarily over the panel area by appropriate selection of the strips. For the number of possible dot positions, the system has fewer electrodes than that shown

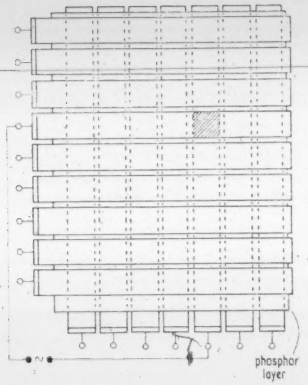


Fig. 4. Electroluminescent display panel with crossed electrode arrays.

on Fig. 2(*c*) but is very much less flexible since only a single spot can have complete freedom of position and movement. Panels of this kind involve a well-known 'cross effect'. It arises from the fact that cross-over points between a particular energized electrode and the remaining unwanted (earthed) electrodes would carry *half* the excitation voltage and would thus emit a certain amount of light. It has been shown that the use of non-linear circuit elements (O'Connel and Narken 1960) as described above can go a long way towards eliminating this effect. In this way, contrast ratios grea.er than 10^4 : 1 have been achieved.

4. Applications of photo-electroluminescence

Photo-electroluminescence is the process whereby the emission intensity of an electroluminescent phosphor is modulated by external radiation. It does not matter fundamentally in this connection whether the original electroluminescence is prominent or not. Cusano (1955), for instance, used thin films (containing manganese and chlorine) deposited by a vapour reaction which showed very little normal electroluminescence, but become very sensitive to electrical stimulus in the presence of x-rays or ultra-violet radiation. There is, as yet, no complete certainty as to how these external radiations affect the electroluminescent process. They could do so by increasing the number of free electrons available for participation or by altering the field conditions within each phosphor grain in such a way as to increase the energy of the colliding electrons. Experiments on single crystals should eventually clarify this issue.

The control by x-rays or ultra-violet radiation is *positive* in the sense that increased radiation results in increased light output. *Negative* control (quenching) can also be achieved, at any rate in principle, by the use of infra-red radiation, as Heckscher (1957) and Narita (1960) have shown. This effect may be due to a diminution of effective field strength (arising from diminished electron occupation of traps) or to a prevention of recombination in certain activator centres (arising from the excitation of valence electrons). The amount of quenching appears to be rather small in the cases so far studied and has not yet found practical applications. Systems are also known in which ultra-violet radiation lowers the light output, whereas x-rays increase it (Kazan and Nicoll 1957). These complicated modes of behaviour are not yet properly understood.

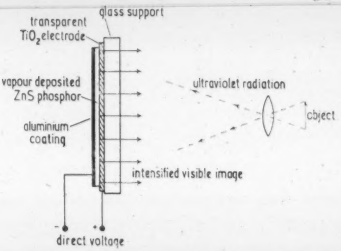


Fig. 5. Image-converter-amplifier based on radiation controlled electroluminescence (*after* Cusano, as described by Kazan and Nicoll 1957).

The radiation effects can be utilized either for providing signal modulation as such or for the purpose of producing a visible picture. A picture display of this kind is shown on Fig. 5. In this case the phosphor film was deposited directly from the vapour. Such films, in contrast to those previously discussed, can show electroluminescence under excitation by d.c. voltages. These devices are, of course, image converters and could produce either positive or negative images. Since electroluminescent phosphors can be made in many different colours, these systems offer a variety of design possibilities, though a large amount of development is still needed. The possibility of producing a negative output image from a positive input is of considerable interest in the context of photographic processing techniques.

It was suggested by Williams (1955) that photo-electroluminescent systems offer the inherent possibility of light amplification, and this was soon experimentally verified. Under favourable conditions, Cusano obtained 10 photons of visible light for every photon of 3650 Å radiation. With increasing irradiation intensity, the amplification ratio diminished. Similar results have been reported by Halstead (1957). An amplifying panel of 4 in. × 4 in. described by Kazan and Nicoll (1957) was capable of displaying and resolving geometrical patterns down to 10 μ in size, though its response time was rather long (of the order of seconds). In principle, there is no reason why amplification should not also be obtained with infra-red radiation. The great drawback of these devices is, of course, that infra-red pictures are available only in rather special circumstances, and ultra-violet pictures very rarely indeed.

5. Devices involving indirect control of electroluminescence

The above forms of control are 'direct' in the sense that the stimuli act upon the electroluminescent process itself. It is alternatively possible to achieve control in a variety of ways with the aid of additional materials or components. Greguss and Weiszburg (1959, 1960), for instance, have described a composite screen system involving an electroluminescent phosphor dispersed in a barium titanate matrix which showed an interesting piezo-effect. The impact of

ultrasonic energy on such a screen while under electrical stimulus resulted in a colour shift of the emission, from the blue–green towards the yellow end of the spectrum. The occurrence of the effect may be critically associated with the presence of the titanate but its mechanism is not yet understood. Leistner (1957) has shown that photoluminescence can be directly (intensity-) modulated by mechanical pressure changes and this suggests that the same may be true for electroluminescence. In one or other of these ways it should eventually become practicable to convert a pressure pattern, whether static or dynamic, into a corresponding visible image (acousto–optical transducer). Modulation by pressure could alternatively be achieved by connecting a pressure sensitive element in series with an electroluminescent panel, either in the form of a distinct component (Fig. 6(*a*)) or (for image

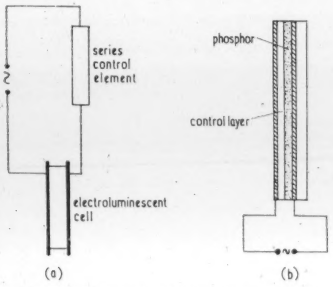


Fig. 6. Electroluminescent devices with indirect control. The series component on (*a*) and the control layer on (*b*) could be piezo-resistive, magneto-resistive, temperature sensitive or photoconductive.

production) in the form of a co-planar layer (Fig. 6(*b*)). By substituting a magneto-resistive component in place of the piezo-resistive one, the system could be controlled magnetically. In a similar way, it could be controlled by temperature if the series element were temperature sensitive. The series component on Fig. 6(*a*) could take the form of a saturable reactor, in which case control could be exercised by a direct current in a secondary circuit. Fig. 9(*c*) shows a variant in which a saturable reactor is used as one of the elements of a resonant circuit. This would make it possible to vary brightness with current in some manner as shown on (*d*), though the curve need not, in practice, be symmetrical.

A technically important method of external control involves the use of series capacitors with ferroelectric dielectrics. Such systems have been proposed (Sack 1959) for use in connection with scanned display panels (see below). The principles are illustrated by Fig. 7, of which (*a*) gives typical characteristics of the series capacitor and (*b*) shows the simplest control circuit. Version (*c*) is actually preferable since it prevents the control voltage from being applied to the electroluminescent cell itself. On this basis, intensity ratios of 100 : 1 have been obtained for control voltages of 200 v. If the series capacitor were entirely free from leakage, then the control would be permanent after the initial charging process. In practice, the control voltage will, of course, decay once its source is removed. The power needed to maintain it permanently is very small indeed. If a bias is superimposed on the control voltage so as to bring the standing brightness about half-way into the useful range, it becomes a simple matter to achieve either 'positive' or 'negative' modulation, depending on the polarity of the video signal.

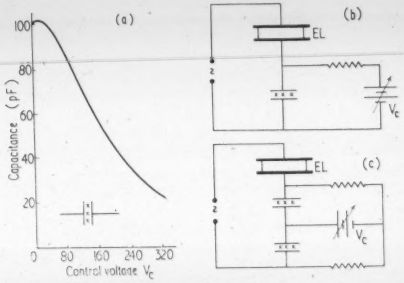


Fig. 7. Control of electroluminescence by means of ferroelectric capacitors (*after* Sack 1959). (*a*) Voltage dependence of typical capacitance; (*b*) simplest control circuit; (*c*) improved control circuit.

Control of a similar kind can be exercised through the use of photoconductive elements. This process is of sufficient importance to warrant separate discussion.

6. Optrons and optron networks

'Optrons' are devices (Loebner 1955) which involve the combination of electroluminescent and photoconductive circuit elements, either by external connection, as here envisaged, or by co-planar construction, as discussed in § 7. They are capable of amplifying electrical and optical signals. Because of the nature of the interaction between the component elements, their spatial relationship (quite unimportant in the cases discussed above) is significant. The coupling can be optical (Fig. 8(*a*)) or electrical (Fig. 6(*a*)) or

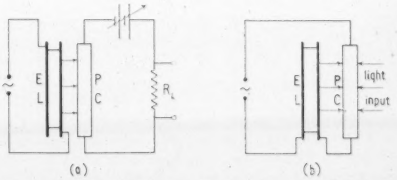


Fig. 8. Combination of electroluminescent (EL) and photoconductive (PC) components (optron networks). (*a*) Optical coupling only; (*b*) optical and electrical coupling; positive feedback.

both Fig. 8(*b*)). The photoconductors are usually based on cadmium sulphide and sometimes on cadmium selenide. The analysis of these constructions is complicated by the fact that both elements are optically and electrically non-linear (Nicoll 1959). Power amplification factors of 100 to 1000 have been reported (Diemer and Van Santen 1960) for the arrangement on Fig. 8(*a*). If the electroluminescent panels were of the kind made by Thornton (1959a) or Cusano (1955), it should be possible to obtain d.c. amplification in the same way. The arrangement (Loebner 1955) on Fig. 8(*b*) provides for positive feedback, always assuming that there is an overlap between the emission spectrum of the electroluminescent cell and the sensitivity spectrum of the photoconductor. With sufficient feedback, the circuit becomes

bi-stable. It can be triggered from the 'off' to the 'on' condition by an increase of the applied voltage or by a light pulse on the photoconductor. Since the brightness (and, indeed, the colour) of the electroluminescent emission is frequency dependent, triggering by a sudden frequency change is also possible if the components are suitably matched. The changeover from the 'off' to the 'on' position is, however, slow because it involves the characteristics of the photoconductor. Loebner has explored the behaviour of such optrons and has found that, with the materials now available, it is necessary to give the photoconductor a certain amount of bias illumination in order to ensure that the critical voltage for bi-stable operation is below the breakdown voltage of the phosphor layer.

It is possible to devise many variants on the present theme. Fig. 9(*a*), for instance, shows an arrangement which combines

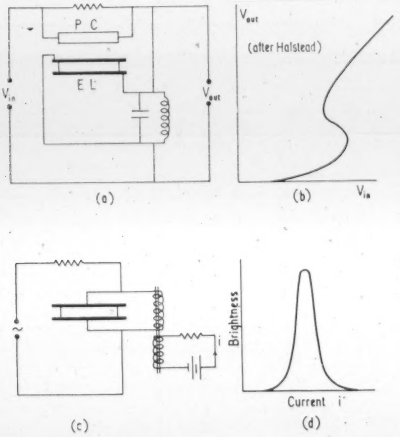


Fig. 9. Electroluminescent elements in resonance circuits.

an optron with a resonance circuit (Halstead 1957). At resonance a high voltage exists across the electroluminescent cell. This reduces the impedance of the photoconductor which, in turn, leads to a higher output voltage. Fig. 9(*b*) gives the corresponding output characteristic. The combination is bi-stable and can again be triggered in three ways. Electrical signal gains of the order of 10^5 have been obtained (Halstead 1957).

The arrangement on Fig. 8(*b*) can give homo- and heterochromatic amplification. *Light* amplification factors of 50 to 100 are easily achieved. It was found possible to match the components very critically so that as little as 10^{-5} of the ultimately emitted flux was capable of triggering the optron. For achieving this condition, there is always an optimum operating voltage. For low voltages the gain is very small and for very high voltages the electroluminescent cell may be in the emitting state even when the photoconductor is in darkness.

Optron networks can perform a variety of logic operations (Loebner 1959), but their slow response rules them out as far as some of the more obvious applications are concerned. Fig. 10, for instance, shows a triple 'or-gate' and a triple 'and-gate'. As a further elaboration, the photoconductive components could be sensitive to different parts of the spectrum, if a particular application were to demand it. By means of optron combinations, it is also possible to devise

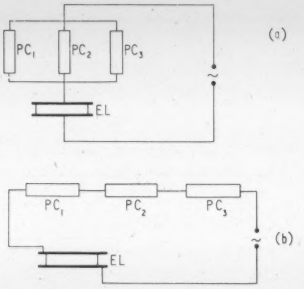


Fig. 10. Simple logic networks employing electroluminescent (EL) and photoconductive (PC) components. (*a*) Triple 'or-gate'; (*b*) triple 'and-gate'.

shift registers (Tomlinson 1957), ring oscillators and flip–flop circuits (Diemer and Van Santen 1960).

7. Amplifying picture screens

If it is desired to amplify a picture rather than a simple light signal, then control of the kind described above must be applied uniformly over the entire picture area. The simplest method (Orthuber and Ullery 1954) is illustrated in Fig. 11(*a*). It is capable of giving picture amplification but suffers from important drawbacks. In order to match the

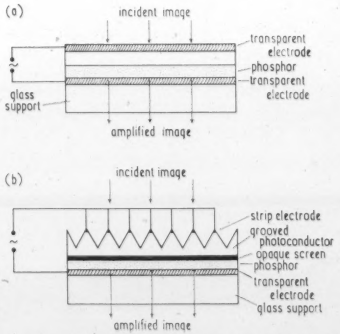


Fig. 11. Amplifying picture screen. (*a*) *After* Orthuber and Ullery (1954); (*b*) *after* Kazan and Nicoll (1955, 1957), with optical feedback prevention.

electroluminescent and photoconductive layers electrically, the photoconductor must have a substantial thickness. However, when this condition is fulfilled, the action of light upon the layer becomes very inefficient. This is so because the incident light is absorbed in a very thin surface film and most of the photoconductor remains unaffected by it altogether. This dilemma led Kazan and Nicoll (1955, 1957) to adopt a different method which is illustrated in Fig. 11(*b*). The grooved photoconductor overcomes the matching difficulty without too great a loss of sensitivity.

In the electrical sense it is a 'thick' layer, but the incident light is able to act upon it in the regions in which it is most effective. In addition, this panel includes an interposed opaque layer which serves to prevent optical feedback. In its absence, light emitted by any part of the phosphor layer would impinge on the neighbouring region of photoconductor. Hence, bright areas would inevitably spread out, resulting first in a distortion of the picture and ultimately in a screen which is uniformly bright. Panels of 12 in. × 12 in. size have been made, giving a homochromatic energy gain of about 100 for the yellow emission band, and acceptable resolution for normal viewing. The applications of such panels could be considerable. One of their first suggested uses would be as x-ray intensifiers for medical radiography. The present limitations arise mostly from the well-known short-comings of photoconductors rather than from the electroluminescent element itself. Kazan (1959) has described a more elaborate system which is a combination of the arrangements shown on Figs 11(*b*) and 12. It has two electroluminescent layers, separated by an opaque film, and provides for a *small* amount of optical feedback between one of these layers and the adjoining photoconductor. In this way, the gain can be increased and the response accelerated without impairing the half-tone picture quality.

8. Light pattern storage panels

In the context of light storage, as opposed to half-tone amplification, close optical coupling between the electroluminescent and photoconductive layers is actually essential. The problem is how to prevent the bright regions from spreading until they cover the whole screen area. This can be done by a cellular construction which can take a number of different forms (Loebner 1959). What is probably the simplest of these is illustrated in Fig. 12. It is effectively a

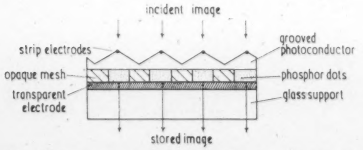


Fig. 12. Light storage matrix. Section through panel (*after* Hook 1959).

matrix of optron units, with cross-coupling prevented by the presence of an opaque mesh. The mesh can be 'photoformed' on glass. Alternatively, it can consist of regularly spaced glass spheres. A 12 in. square panel with 250 000 storage cells has been described by Hook (1959). When such a panel is 'shown' a black-and-white pattern, it will duplicate and store it.*

The information can be 'wiped off' by removal of the external voltage. Functionally, at any rate, we have here another link between the technologies of electroluminescence and photography.

9. Scanned picture screens with ferroelectric and transfluxor control

The panels discussed in § 7 all require an optical input picture. In principle, pictures can be formed by the use of electrical video inputs alone. This most desirable of aims (which could lead to a virtually two-dimensional television screen) is also technically the most difficult. The two principal requirements are, of course, to provide for the possibility of scanning and for intensity modulation. In addition, it would be necessary to consider problems of resolution and response time. Unfortunately, the simplest scanning system illustrated on Fig. 4 is useless in this connection. If the panel were to consist of only 10^4 dots and the frame repetition rate were 25 per second, then each dot would be energized for 4 μs only. Using an audio frequency supply as usual, this would amount to a fraction of one cycle. In these circumstances the emission would be very small, and indeed visually undetectable. The impossibility of superimposing an intensity modulation is also obvious. Three quite different systems have therefore been explored. Two of these consist of dot electrode systems as on Fig. 2(*c*), with external control elements connected to each dot. The 'Elf-screen' described by Sack (1959) employs control by means of ferroelectric capacitors, as shown on Fig. 7. The repetition of this structure for each dot of a picture panel is, of course, technically difficult. The ferroelectric material is first deposited as a uniform layer and then machined into the matrix form. The matrix is thus an integral part of the screen. The distribution of the scanning signal is an equally complex problem but a number of suggestions have been made as to how this might be done. Performance figures of the device as a whole are not yet available.

A second type of screen employs the same kind of electrode system but uses magnetic 'transfluxors' for control purposes in place of ferroelectric capacitors, each picture dot being controlled by its own transfluxor. The transfluxors themselves are made of ferrite material and, in spite of their complexity, can be very small. (0·08 in. outside diameter has been quoted.) Their operation has been described by Rajchman and Lo (1955, 1956). The arrangement (Rajchman *et al.* 1958) for a single dot is shown on Fig. 13(*a*). For present purposes, it is sufficient to regard the transfluxor as a special kind of transformer. The electroluminescent element is connected to the secondary winding, and the primary carries a pulsed drive signal. The amount of coupling between primary and secondary depends on the state of magnetization of the ferrite core and this can be controlled by sending currents through the X and Y leads. The setting can be achieved by a single pulse and, once made, is permanent until a new setting pulse alters the situation. Such a pulse can saturate the core completely, in which case the coupling between primary and secondary will be zero and the electroluminescent element will remain dark. Alternatively, a pulse (through either X or Y leads) could saturate the core partially (i.e. completely up to a certain radius r and not at all for larger distances), in which case the coupling would be intermediate between zero and the maximum value possible. The detailed setting characteristics are such that only the simultaneous presence of current pulses in the X and Y leads is sufficient to produce an appreciable output flux and hence light emission from the electroluminescent cell. This is apparent from a consideration of Fig. 13(*b*). The coincidence principle is, of course, the basis of the scanning procedure. Each X lead is linked with a whole row of transfluxors, each Y lead with a column of identical

* An interesting storage panel has been developed by Ranby, Hobbs and Turner in which all the above functions are somehow combined in a single layer. Constructional details are not yet available. The device operates from a d.c. source, and provides good half-tone rendering, though the information is stored satisfactorily for a short period only, e.g. about a minute (British Patent Appl. No. 26957/59). I am indebted to Dr. S. T. Henderson of Thorn Electrical Industries Ltd., for a demonstration of the device.

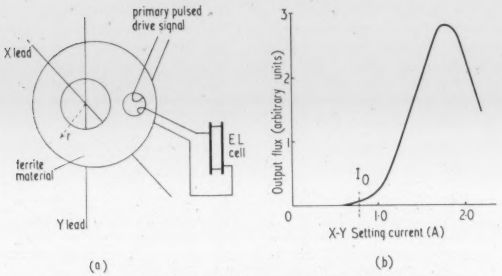


Fig. 13. Operation of an electroluminescent element with transfluxor control. (*a*) Connections to transfluxor; (*b*) setting characteristic (*after* Rajchman, Briggs and Lo 1958).

units. One set of leads could carry constant pulses of magnitude I_0; the other set would carry variable pulses corresponding to the video information. The device could thus operate mostly within the linear region of the setting characteristic.

Rajchman, Briggs and Lo (1958) have described the characteristics of a screen constructed in this way. The experimental model consisted of 1200 dots, a large number considering the difficulties involved, but too small a number to give good resolution. The frame rate was only 15 per second, but this gave a good illusion of continuous action, because each element was in the emitting state continuously for 29/30 of the frame time. The screen had good half-tone characteristics and adequate brightness for normal viewing.

10. Scanned picture screens with piezovoltaic control

Any display screen which consists of a dot matrix and individual control elements is necessarily costly. Moreover, as the screen size increases, uniform structure becomes very difficult to achieve. In an attempt to overcome these shortcomings and to simplify the electronic drive mechanism, an entirely different form of approach has been proposed by Yando (lecture delivered at the International I.R.E. Convention, New York, March 1961, and private communication). It depends on a combination of electroluminescent and piezoelectric layers. The principle and design are illustrated on Fig. 14. Fig. 14(*a*) shows a thin piezoelectric panel, made of lead zirconate–titanate which is a ceramic material. One of its surfaces is entirely covered with a metallic base electrode, the other carries two narrow strip electrodes as shown. The application of a voltage pulse between base and a strip leads to the formation of an elastic wave which is propagated through the panel almost without loss and is eventually absorbed (without reflection) in the acoustic termination. The elastic wave is accompanied by a localized piezoelectric voltage, developed across the thickness of the panel. Voltage pulses in straight line form can thus be made to travel across the panel in two directions, emanating from each of the strip electrodes. 3600 m/s has been quoted as a typical velocity. At their cross-over point the two waves would reinforce each other and the voltage developed would therefore be twice as high. If simultaneous elastic waves were to be produced by each of the strip electrodes, then it is clear that the cross-over point would travel across the sheet along a diagonal line. By suitable phasing of the two pulses, this line can be displaced and a complete scanning pattern can thus be achieved.

Figure 14(*b*) shows the construction of the panel as a whole. A non-linear resistance layer is included for the

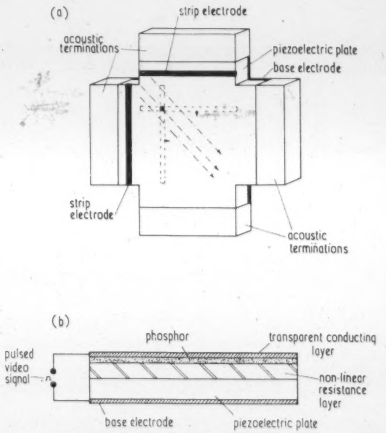


Fig. 14. Scanned picture screen with piezovoltaic control (*after* Yando) (lecture delivered at the International I.R.E. Convention, New York, March 1961 and private communication). (*a*) Operating principle; generation of the elastic wave pattern; (*b*) screen structure and method of superimposing video signal.

purpose already outlined in § 3 above. The layers are matched so that the double voltage generated at a cross-over is only just sufficient to produce an electroluminescent signal, and the single voltage generated elsewhere is not. The scanning process thus produces a uniform field of luminescence, presumably near the threshold of visibility, on which positive video modulation can then be superimposed. The piezoelectric voltage and the (smaller and pulsed) video signal are effectively connected in series and the diagram shows how this is done. Yando has made and demonstrated experimental panels of this kind up to 5 in. × 5 in. in size. With the appropriate electronic controls, they are capable of displaying normal oscilloscope patterns, with a brightness of 0·1 ft-lambert and an effective spot size of about 1 mm. The device is still in the development stage. Its simplicity of construction and the ease with which the video information

can be distributed probably make it the most promising solid state display for television purposes at the present time.

11. Future developments

The distant future is far from clear, but it is reasonable to believe that interesting developments along the above lines will continue during the next few years. The greatest research needs are not, however, concerned with structural and configurational problems but with the understanding of electroluminescence itself. Accidental discoveries apart, it is only in this way that one may hope to increase the efficiency and the operating life of electroluminescent layers.

In the most widely held view, light emission from the familiar powder phosphors involves the impact ionization of activator centres under the influence of the effective field. This is essentially an inefficient process because only a fraction of the electron collisions could ever be energetic enough to produce ionization. This fraction should, however, depend on the mean free path of the carriers, and one may therefore hope to increase it by using crystallites with a high degree of short-range order. This would eliminate harmful defects and may also have the secondary consequence of reducing unwanted *non*-radiative recombinations. However, quantitative experiments along such lines have not yet been practicable. They would be critically dependent on advances in our preparational techniques and, indeed, on our means of assessing short-range disorder. A further problem is the stabilization of activator centres in the presence of high fields. Under such conditions, ionic motion is always likely. As yet, we do not know enough about the manner in which activator centres are accommodated to devise ways and means of inhibiting this motion.

There is another possibility which warrants careful exploration. By means of a vacuum evaporation process, Thornton (1959b) has made a number of phosphor films which have shown electroluminescence at excitation levels so low (e.g. 2·2 v) that ordinary impact ionization must be ruled out. The interpretation of this effect could evidently have a bearing on future developments. We are dealing here with an electroluminescent equivalent of the phenomenon which is known as 'anti-Stokes emission' in the context of photoluminescence. Pátek (1961) has tentatively suggested that it could arise from a phonon-aided impact excitation process. However, such a mechanism would be enormously temperature dependent in a manner which is opposite to that actually observed. The interpretation is therefore not acceptable. If, on the other hand, low-voltage electroluminescence were to be due to minority carrier injection, as Thornton has proposed, or due to carrier accumulation, a possibility envisaged by Henisch and Marathe (1960), then the prospect of increasing the efficiency and of eliminating the need for high fields would exist in principle. At any rate, the operational limitations would then be different. The brightness levels which can be achieved by small excitation voltages are always low, as would be expected. However, films of the same kind have been tested (Thornton 1961) at higher voltages and have been found capable of yielding brightness levels of quite unusual magnitude. Thus, a yellow emitting zinc sulphide film, containing copper, manganese and chlorine, gave electroluminescence at 600 ft-lamberts when excited by 85 v d.c. There is, as yet, no information on its life expectancy. It is clear that thin films offer attractive opportunities for future development, even though much systematic experimentation will be needed before the operative mechanism is unambiguously resolved. There is also a great deal of scope for the exploration of crystals other than zinc sulphide and zinc selenide, partly in the hope of finding new materials of technological value and partly in order to broaden and to generalize the work of theoretical interpretation.

Acknowledgments

I should like to thank P. C. Banbury, E. A. Faulkner and E. A. Davis of this Department, for their helpful suggestions, and S. Yando of General Telephone and Electronics Laboratories Inc., and W. A. Thornton of the Westinghouse Electric Corporation, for generously providing unpublished information.

References

Cusano, D. A., 1955, *Phys. Rev.*, **98,** 546.
Diemer, G., and Van Santen, J. G., 1960, *Philips Res. Rep.*, **15,** 368.
Gillson, J. L., 1956, U.S. Patent 2,733,367.
Greguss, P., and Weiszburg, J., 1959, *Acustica*, **9,** 183.
—— 1960, *Acta Phys. Hung.*, **11,** 185.
Halsted, R. E., 1957, *Proceedings of the Symposium on the Role of Solid State Phenomena in Electric Circuits* (Polytechnic Institute of Brooklyn).
Heckscher, H., 1957, *J. Opt. Soc. Amer.*, **47,** 765.
Henisch, H. K., and Marathe, B. R., 1960, *Proc. Phys. Soc.*, **76,** 782.
Hook, H. O., 1959, *R.C.A. Rev.*, **20,** 744.
Ivey, H. F., 1959, *I.R.E. Transactions of Prof. Group on Electron Devices, ED-6*, No. 2.
—— 1961, *J. Electrochem. Soc.*, **108,** 590.
Kazan, B., 1959, *Proc. Inst. Radio Engrs, N.Y.*, **47,** 12.
Kazan, B., and Nicoll, F. H., 1955, *Proc. Inst. Radio Engrs, N.Y.*, **43,** 1888.
—— 1957, *J. Opt. Soc. Amer.*, **47,** 887.
Larach, S., and Shrader, R. E., 1956, *Phys. Rev.*, **102,** 582.
Leistner, M., 1957, *Ann. Phys.*, **20,** 129.
Loebner, E. E., 1955, *Proc. Inst. Radio Engrs, N.Y.*, **43,** 1897.
—— 1959, *R.C.A. Rev.*, **20,** 715.
Lossew, D. W., 1923, *Telegrafia i Telefonia*, **18,** 61.
Mager, E. L., 1951, U.S. Patent No. 2, 566, 349.
Mash, D. H., 1960, *J. Sci. Instrum.*, **37,** 47.
Narita, S., 1960, *J. Phys. Soc. Japan*, **15,** 128.
Nicoll, F. H., 1959, *R.C.A. Rev.*, **20,** 658.
O'Connel, J. A., and Narken, B., 1960, *I.B.M. Journal of Res. and Dev.*, **4,** 426.
Orthuber, R. K., and Ullery, L. R., 1954, *J. Opt. Soc. Amer.*, **44,** 297.
Pátek, K., 1961, *Czech. J. Phys.* B, **11,** 18.
Rajchman, J. A., and Lo, A. W., 1955, *R.C.A. Rev.*, **16,** 303.
—— 1956, *Proc. Inst. Radio Engrs, N.Y.*, **44,** 321.
Rajchman, J. A., Briggs, G. R., and Lo, A. W., 1958, *Proc. Inst. Radio Engrs, N.Y.*, **46,** 1808.
Rulon, R. M., 1955, *Sylvan. Tech.*, **8,** 45.
Sack, E. A., 1959, *Research, London*, **12,** 54.
Thornton, W. A., 1959a, *Phys. Rev.*, **113,** 1187.
—— 1959b, *Phys. Rev.*, **116,** 893.
—— 1961, *Phys. Rev.*, **122,** 58.
Tomlinson, T. B., 1957, *J. Brit. Instn Radio Engrs*, **17,** 141.
Williams, F. E., 1955, *Phys. Rev.*, **98,** 547.
Zalm, P., Diemer, G., and Klasens, H. A., 1955, *Philips Res. Rep.*, **10,** 205.

Vacuum-deposited metal film resistors

by G. SIDDALL and B. A. PROBYN, Research Laboratories, Edwards High Vacuum Limited, Crawley, Sussex

MS. received 10*th July* 1961

Abstract

Although vacuum-deposited thin metal films have been studied for many years, it is only recently that their anomalous electrical properties have been turned to practical account. In this work the most important requirements for the deposition of metal film resistors are discussed, and an annealing cycle up to 300–350° C *in air is proposed to stabilize the resistance of nickel–chromium alloy films, preventing further large irreversible change within the temperature range of normal use. The stability of annealed unprotected films within the range* 10–400 Ω/*square was determined over a period of* 2000 *h on load* (1 W/in^2) *and* 8000 *h at no load. Frequency diagrams show the number of resistors measured, and their percentage resistance change. Resistors at no load increased by an average of* 1% *compared with a very small mean decrease in the same resistors on load. The temperature coefficient of resistance of films in the range* 100–150 Ω/*square is shown to vary between* 0·002 *and* 0·1% *deg*$^{-1}$C *and is related to the rate of film deposition. Starting with an alloy of* 80/20 *nickel–chromium, the lowest values of temperature coefficient of resistance are obtained for rates of deposition lower than* 5 Ås^{-1} *and this effect is discussed in terms of the oxidation of the film and the separation of the alloy during evaporation.*

1. Introduction

At the beginning of this century, Thomson (1901) proposed a theory to explain why the resistivity of thin metal films should be greater than that of the bulk metal. He deduced that the mean path of the charge carriers (later to be called the free electrons) between successive interactions or collisions with the atoms in a thin conductor was shortened by the closeness of the boundaries. The resistivity was increased because of the energy lost at each collision. This idea was extended by Planck (1914), Fuchs (1938), Lovell and Appleyard (1937) and others, who tried to relate the increase of resistance to the decrease of thickness. Unfortunately many of these attempts were based upon the assumption of an ideal thin film with parallel boundaries, free from defects and regular in structure. In view of present-day knowledge available from numerous electron microscope studies of film structure, it is not surprising that the above work was unable to explain the very wide variation of experimental results. It is now known that films deposited by evaporation in a vacuum are usually highly imperfect, and this fact is sufficient reason for the resistivity to become much higher than the value of bulk resistivity for the deposited metal.

Recent experimental work has shown that under carefully controlled conditions, ordered films can be deposited on surfaces of regular crystal structure. For example, Bassett and Pashley (1961) have succeeded in growing oriented gold films on cleaved rocksalt and on silver surfaces. Preston, Gillham and Williams (1955) have deposited gold films of high electrical conductivity by using a glass substrate previously coated with a thin layer of bismuth oxide. Such films have a resistivity which closely approaches the bulk value, but considerable crystal imperfections, e.g. dislocations and stacking faults, are still present. It is thus unlikely that the boundary effect contributes to the high resistivity of very thin metal films as significantly as earlier workers had thought.

Although the resistivity of films is higher than that of the bulk metal, the use of thin films for resistance elements is limited by a fundamental problem. Metal films having a useful resistance value are usually unstable. This effect has been studied by Vand (1943) who has shown that the irreversible decreases of resistance, which occur upon heating a thin film, can be explained by re-ordering of the crystal lattice due to a 'decay of the lattice defects'. These changes are much larger than resistance changes due to the normal (positive) temperature coefficient of resistance of the material. Vand has proposed that a new metal film has lattice disorder with vacant sites and excess atoms situated so near that the unstable equilibrium is easily disturbed. Separate energy levels exist, corresponding to different types of defect in the structure, and a definite temperature is required to remove the defect at each energy level. The resistance of the film falls to an irreversible value, and this structure should be stable at temperatures below its previous annealing point. This stability is not realized in practice because of oxidation or absorption of atmospheric water vapour, but stability can be improved by suitable protection.

Vand's work was based upon the annealing of films deposited at liquid air temperatures. When films are deposited at about 20° C or higher, the decreases of resistance observed during subsequent heating are less marked, and tend to be compensated by resistance changes due to further oxidation, agglomeration or recrystallization.

After annealing, the resistivity of a thin metal film remains higher than the bulk value, but the temperature coefficient of resistance is usually lower. It has been shown by Holland and Siddall (1953) that the increased resistivity of gold films within the range 40–400 Å can be related to the decrease of temperature coefficient of resistance. The contributions to the resistivity from lattice imperfections, boundary effect and gas inclusion are greater in a thin film and, over a limited temperature range (after annealing), are reasonably independent of temperature. Thus the temperature-dependent part of the resistivity is comparatively less than in the bulk material.

It has also been observed that the temperature coefficient of resistance of a very thin conductor is often negative. For example, thin nickel and iron films deposited by cathodic sputtering were shown by Itterbeeck *et al.* (1951) to possess a high negative temperature coefficient of resistance. Holland

and Siddall (1953) have pointed out that such films were deposited under conditions in which oxidation was likely to occur; thus the observed properties could be attributed to the formation of a partially oxidized metal layer behaving as a semiconductor.

The erratic behaviour of thin metal films is well known and is the subject of an extensive literature, but as shown by the foregoing, a better understanding of film properties is beginning. Summarizing the Introduction, it can be stated that the anomalous electrical properties of films are principally due to their structural imperfections and to the thermodynamic instability produced when metal vapour is abruptly condensed to the solid phase. The changes of resistivity and temperature coefficient of resistance, which occur upon heating or ageing a film, arise from the re-ordering of the structure, the relief of high internal stresses and the further oxidation or gas absorption of the film. These changes are parallel to those occurring in fine resistance wires upon annealing after cold-working. However, the gas absorption and higher degree of lattice imperfection in vacuum-deposited films cause much greater variation of properties. There is an increasing amount of evidence that high stability resistance films can be obtained by correct annealing treatment and suitable protection. It is the purpose of this paper to describe a method of making reasonably stable resistance elements by the vacuum deposition of nickel–chromium alloy on glass and to discuss their properties in terms of the processing conditions.

2. Practical requirements

Substrate

The surface of the supporting substrate should be smooth and uniform, and both chemically and mechanically stable at temperatures up to about 350° C in atmosphere and vacuum. Any variation of surface smoothness gives a corresponding variation of film resistance value, because the film is thin enough to be greatly affected by the state of the surface. For example, a film of resistance as low as 10 Ω/square on a polished glass surface may be discontinuous when deposited under identical conditions on a finely ground glass surface.

It is a characteristic of a film deposited from the vapour that the grains tend to grow on surface prominences, which trap the atoms first arriving there and act as centres for nucleation. Films as thick as 1000 Å may be discontinuous when deposited on a coarse surface because large grains are formed which do not touch, and the thickness must be increased before conductivity is observed. Such films tend to be unstable because their conductivity depends upon contacts between large grains.

The most suitable substrate materials are found amongst glasses and ceramics. Good results have already been obtained using glasses of high silica content such as Pyrex or Vycor. These are two of the few glasses unaffected by water vapour. Many other glasses, including some borosilicates, devitrify in contact with water, and their surfaces become powdery because small crystals of metal silicates are formed. Soda-lime glasses are not used because their surfaces are also chemically unstable. During flame polishing, when the glass is bombarded by thermally produced gas ions or ionic bombardment in a glow discharge, free sodium ions are active at the surface of the glass. They combine with water vapour from the gas atmosphere to form sodium hydroxide and deliquescent sodium silicate by reaction with silica in the glass. Some further reaction with the deposited film must be expected.

Ceramics possessing good chemical and mechanical properties are available; however, their surface smoothness is often variable because of the sintering process used in their manufacture. Glazing is not always a satisfactory solution to this problem because standard glazes are often based upon some of the unsuitable glasses already described. Very careful examination of surface smoothness is needed when choosing a ceramic material for the support of vacuum-deposited films.

The temperature coefficient of linear expansion of metals is usually an order higher than that of glass or ceramics, and this factor partly contributes to the high internal stresses which have been observed in thin films. However, once the films have been annealed, the effect of the expansion of the base on the resistance of the film is very small, compared with the average temperature coefficient of resistance of films, and is insignificant when compared with that of bulk metals.

The resistance alloy

In the early stages of deposition of a metal film, aggregates of metal atoms (nuclei) are formed on the substrate. The number of nuclei is dependent upon the physical and chemical properties of the metal and substrate, and upon the rate of deposition. As the nuclei increase in size they grow together, eventually to form a continuous film. The second stage of growth is marked by the onset of electrical conductivity, and the rate of change of resistance with film thickness is very high. Unfortunately, the most useful resistance values coincide with this unstable region of thickness for many metallic conductors.

The most successful high-resistance films have been made by depositing chromium and alloys of chromium with nickel, silicon, titanium, etc. For example, nickel–chromium alloys have a high bulk resistivity (80–130 $\mu\Omega$ cm) and therefore films of this alloy are much thicker than films of the pure metals for the same resistance value. Films of resistance 400 Ω/square are at least 80 Å thick, more or less continuous and are outside the very unstable region of thickness. Nickel–chromium alloys also have a low temperature coefficient of resistance in bulk, and are very resistant to chemical attack because of the compact protective oxide layer which forms in contact with an oxidizing atmosphere. The formation of double oxides having a spinel structure has been shown on nickel–chromium alloys under examination by electron diffraction, and this reason has been given to account for their high chemical stability.

Evaporation conditions

The lowest values of temperature coefficient of resistance and the best stability are achieved in films deposited under conditions favouring oxidation. During deposition the substrate is heated to relieve the internal stress in the film, but this treatment can also increase the rate of oxidation. The residual gas atmosphere in the chamber of a kinetic vacuum system is highly oxidizing due to the high proportion of water vapour at the normal working pressure (10^{-4} mm Hg). Assuming that the partial pressure of water vapour is only 10^{-5} mm Hg, then it is calculated approximately that 5×10^{15} molecules $cm^{-2} s^{-1}$ strike the substrate surface. If the rate of deposition of chromium metal is about 3 Å s^{-1}, then ten water vapour molecules strike the surface for every

chromium incident atom. Thus there is sufficient oxygen (in the form of water vapour) available at the source for the film to be highly oxidized at normal rates of deposition. Oxidation also occurs after the chromium atoms have left the vapour source, giving rise to the familiar gettering effect. The fall in pressure can readily be observed on the vacuum gauge.

Thus the first few atomic layers deposited during the gettering period are highly oxidized, and when the chamber has been 'cleaned up' the deposit is more metallic. After evaporation ceases, the deposited film remains open to oxidation. Thus the deposited film is inhomogeneous and approximates to a sandwich layer of oxide/metal/oxide, in which the two outer layers are more highly oxidized than the inner layer.

The exact state of oxidation of the deposited film is unknown and a further effect of oxidation can be observed upon baking in air. The final resistance change upon annealing may then be positive or negative, because the decrease attributed by Vand to lattice transformation may be greater or less than the increase due to further oxidation.

Heat treatment and protection

Heat treatment carried out during or after deposition serves three purposes:

(i) high internal stresses in the film are relieved;
(ii) some defects in the crystal lattice are removed, thus improving the heat-stability;
(iii) a protective oxide layer is completed, making the film less subject to external atmospheric attack.

In practice, it has been found advisable to heat the substrate in vacuum before deposition to a temperature of at least 300° C. A further heating period in air for 30 min at 300° C completes the annealing of the film.

The electrical properties of resistance tapes and wires are stabilized by annealing and by cyclic baking in air or hydrogen. This treatment reduces the strains and dislocations set up during the drawing of the wire. Thus the treatment required by a vacuum-deposited film is similar. Baking during and after deposition re-orders the crystal lattice, and improves the resistance stability with time, also forming a compact oxide surface layer. Several fast baking cycles carried out in air hasten the changes of resistance up to 300° C, which become smaller with each successive cycle.

3. Experimental work

Evaporation technique

The preparation of nickel–chromium resistance films was carried out in a vacuum deposition plant having a 12 in. diameter chamber equipped with pumps capable of reducing the residual gas pressure in the vacuum chamber below 10^{-4} mm Hg in 5 minutes. Provision was made for two h.t. lead-through electrodes (for cleaning by positive ion bombardment), three electrodes for the evaporation source and several smaller electrodes for connecting the radiant heater, thermocouple, and resistance monitor.

The evaporation source was heated by electron bombardment (Fig. 1). This source consisted of a stainless steel supporting block (forming the anode) on which was mounted a $\frac{1}{4}$ in. diameter special ceramic hearth $\frac{1}{4}$ in. high. Nickel–chromium wire (22 s.w.g.) was fed through a stainless steel guide tube to the centre of the hearth. The feed mechanism was mounted at the side of the hearth, and allowed the wire to be fed either continuously or to be intermittently operated by a handwheel outside the vacuum chamber. The nickel–chromium wire was bombarded by electrons emitted from a small hot filament of 0·020 in. diameter tungsten

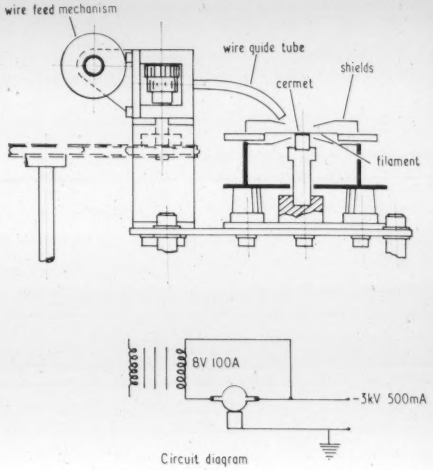


Fig. 1. Schematic diagram of the evaporation source heated by electron bombardment.

wire (forming the cathode), supported $\frac{1}{8}$ in. above the top of the hearth. The cathode heater supply was obtained from a 8 v, 100 A transformer with secondary winding insulated from earth and primary for 15 kv. The anode and cathode were connected across a suitable h.t. supply, having a maximum power of 1·5 kw at 3 kv; the anode was held at earth potential and the cathode at negative 3 kv. The top of the hearth was hollowed to enable the wire to melt and form a bead from which evaporation could take place.

Substrates and workholders

Special jigs were made to hold flat specimen plates of Pyrex, soda glass and ceramic. During each evaporation, the resistance of one plate was monitored by connecting the end terminals to an external circuit for resistance measurement. A simple ohmmeter was used for monitoring the resistance value during evaporation. The accuracy of the measurements was only of the order $\pm 2\%$, but the results were used only to indicate the approximate value of the resistance during evaporation. A bridge method of measurement was used to determine accurately the resistances of the slides, and is described more fully later.

The workholder consisted of a simple jig constructed so that only $\frac{1}{8}$ in. at each end of the slides was masked by the clamp, and these were placed next to the monitor plate. The workholder was supported 4 in. above the evaporation source by means of a tripod. A radiant heater, dissipating 750 w at 110 v, was mounted above the workholder to raise the temperature of the substrate to 300° C before evaporation. The temperature of the substrate was measured by means of a chromel–alumel thermocouple placed inside the vacuum chamber with its junction resting on the top face of the workholder. The thermocouple was connected to an external

meter calibrated to read degrees Centigrade, covering a temperature range from 0 to 500 in divisions of 10 degrees.

A special chamber assembly was constructed for deposition of films on cylindrical formers and, for monitoring their resistance, a static flat glass slide was used. By experiment a simple relationship between the resistance of the static plate and the resistance of the cylindrical formers was obtained, thus enabling the evaporation to be roughly monitored.

Contacts

The method of making contact to the deposited film influenced both the accuracy with which the film could be measured and the ultimate stability. To reduce the contact resistance to a negligible value, thick nickel–chromium films (1500 Å) were deposited at the ends of the slides in a separate vacuum deposition cycle. The contact between the thin resistance film and the thick terminal was of large area compared with the cross section of the film. Copper leads were soldered directly to the thick nickel–chromium terminal using an indium solder without flux.

Deposition procedure

The slides were thoroughly cleaned with a detergent, washed in water and *iso*propyl alcohol, and finally degreased in *iso*propyl alcohol vapour, before being placed in the work chamber. The chamber was evacuated to a pressure of less than 10^{-3} mm Hg and the substrates heated to a temperature of 250° c by means of a radiant heater. The slides were then h.t. cleaned for a period of 10 min, after which time they were further heated by the radiant heater to a temperature of between 300° c and 330° c.

The pressure in the chamber was adjusted to the required value and nickel–chromium was deposited on the substrates for a measured period of time until the correct resistance value was reached. The samples were allowed to cool to below 100° c before removing them from the chamber.

Stabilization

Stabilization or annealing of the resistance films was carried out in a special oven in air at atmospheric pressure. The oven consisted of a cylindrical steel body 5 in. diameter × 7 in. long, around which was wound a resistance heater. One end of the cylinder was sealed, with provision made for inserting a thermocouple to measure the temperature of the film during baking. A special jig, capable of holding several resistors, was placed in the unsealed end of the oven. Each resistor was clamped individually using pure nickel foil to make good contact, and the clamping screws were connected to a selector switch so that the resistance of each element could be measured separately during the heating cycle.

The temperature of the oven was raised to 300° c in about 20 minutes and maintained at this level for 30 minutes. The oven was then allowed to cool before the resistors were removed.

Readings of the individual resistance values against temperature were taken during the whole of the stabilizing cycle. Final resistance measurements were taken when the specimens had cooled to room temperature.

Measuring apparatus

Resistances were measured by means of a Wheatstone bridge circuit and the accuracy of measurement was estimated to be of the order of 0·1%.

For measurement of the temperature coefficient of resistance, the resistors were totally immersed in a silicone-oil bath which was heated by means of a flat resistance heater. The liquid was stirred mechanically to ensure an even temperature distribution; the resistor under test was connected to the Wheatstone bridge and a mercury thermometer was clipped to one end of the resistor. Changes in resistance were noted over a temperature range from room temperature to 110° c for every increment of 5 deg c. These measurements were taken both on heating-up and cooling-down of the liquid. Average values of resistance were taken for the corresponding temperature, and resistance was plotted against temperature. From this the temperature coefficient of resistance was calculated for each element.

One slide in each batch was partially masked by a very thin steel strip to obtain a sharp step from the film to the glass. This glass was located next to the monitor slide, and after the nickel–chromium film had been deposited, the mask was removed, and the thickness of the film was determined by multiple beam interferometry (Tolansky 1948).

Stability

Stability tests were carried out on 52 flat resistors within the range 10–400 Ω/square selected for good contact (after annealing) for a period of fifteen months, three months under load with each slide dissipating 3 w, and twelve months at no load. The slides were mounted on a 'peg board' by means of clips. The clips were connected to tappings on an autotransformer adjusted so that the correct voltages were applied across each resistor to give a power dissipation of 1 w/in^2. During the load tests, resistance measurements were carried out daily for the first week, and then every week for the remainder of the three-month period. The peg board and slides were protected from dust and dirt by a polythene sheet. After the load tests were completed, the sheet was removed and the slides left on shelf-life for the next twelve months with the power switched off. During this period the resistance values were checked weekly and the results tabulated. When the stability tests had been completed, the total percentage change of resistance of each plate was calculated.

4. Results and discussion

Comparison of substrate materials

The resistances of films deposited in the same batch on Pyrex and soda-lime glass and on a special grade of sintered aluminium oxide were compared in a total of over 50 batches. No difference could be observed between the two glasses in initial resistance value (within the normal experimental variations), but the values of resistance of films on the alumina base were frequently several times higher. Subsequent annealing showed that the ceramic base was unsuitable because the films could not be stabilized, the resistance value increasing with each successive annealing cycle. Some films deposited on soda-lime glass showed a curious effect upon applying a low d.c. potential (about 4·5 v) for monitoring the resistance during deposition at 350° c. When the films were examined, a thin line corresponding to a lower optical absorption was noticed along the negative terminal. This was thought to be due to a reaction of the film with sodium ions transported to the negative terminal of the monitor slide, but microscopic examination revealed nothing but a change of optical density. However, this phenomenon was considered to be of sufficient importance to reject the use of soda-lime glass for resistors, on the assumption that a d.c. field and sufficiently high temperature for the transport of sodium ions might be encountered in normal use. It is

intended to study this effect in more detail. The remainder of the results refer only to films deposited on Pyrex glass.

Resistance changes during annealing

In order to obtain results that could justifiably be compared, these measurements were carried out on over 100 films having the same order of resistance value (in the range 100–150 Ω/square). When the films were baked in air their resistance changes were measured and compared in relationship to the evaporation conditions. The largest permanent decrease of resistance recorded during the annealing of films at 300° c was 8%, but resistance increases up to 25% were obtained. Generally, however, the irreversible resistance changes were within the limits −5%, +10%.

It was established that for the same pressure conditions, slowly deposited films (1 Å s^{-1}) decreased in resistance after baking, whereas films deposited at 10 Å s^{-1} increased on baking. It was also found that if the working pressure was increased from 5×10^{-5} mm Hg to 5×10^{-4} mm Hg by admitting air, then a rapidly deposited film (10 Å s^{-1}) also showed a decrease in resistance after baking. Thus the resistance of a film deposited under conditions favouring gas absorption generally decreased on baking. If the surface layer of oxide did not form completely during deposition, then the film was further oxidized during the baking process, often giving rise to an increase of resistance. Very thin films of high resistances (above 2000 Ω/square) were completely oxidized when baked in air at 300° c and became virtually non-conducting.

It was also found that the order of resistance change during annealing to be expected from a given set of evaporation conditions could not be predicted, but the direction of change, i.e. positive or negative, was more definite. The maximum variation of resistance value in any one batch of resistors was of the order 20%, but this was due in part to differences of physical dimensions. The irreversible resistance changes occurring upon annealing a complete batch were also variable in value, within the limits −5%, +10%, but the majority of resistors in a batch followed the same direction of change. A few anomalies were observed, and could only be explained by bad contact or by a reaction between film and glass surface.

After a second annealing treatment the irreversible resistance changes were in most cases less than 0·5%. It was found that the resistance changes up to about 150° c were reversible for films annealed at 300° c. Thus for most practical purposes, one stabilizing treatment is sufficient; however, the high temperature stability is improved by further cyclic treatment.

The variable changes of resistance upon annealing and the spread of resistance values within a batch were to be expected, because of the possible variations in the flame-polished Pyrex surface and of the vacuum conditions. It is intended in future experiments to study the resistance changes in films deposited at ultra-high vacuum, and to obtain a mass spectrogram of the residual gas atmosphere during evaporation, in order to obtain a better experimental foundation for analysis.

Stability

The results for 51 resistors within the range 10–400 Ω/square have been presented in the following manner. The percentage change in value of each resistor has been calculated to the nearest 0·1% after 2000 h (nominally three months) on load at 1 w/in². The number of resistors possessing the same resistance change has been plotted in the frequency diagram (Fig. 2). It can be seen that there is a slight total overall decrease of resistance.

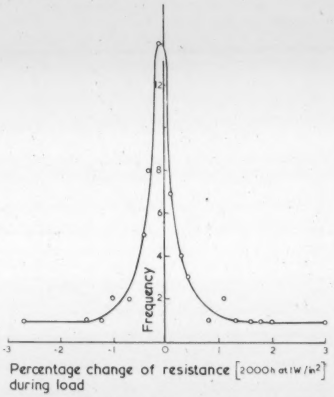


Fig. 2. Frequency diagram, the ordinate at each point corresponds to the number of resistors having the same percentage resistance change (to the nearest 0·1%) after 2000 h dissipating 1 w/in² of film area.

The same method has been adopted for showing the stability of the same resistors at no load (Fig. 3) over a period of 8000 h (nominally twelve months). The frequency curve shows an overall increase of resistance amounting to about 1%.

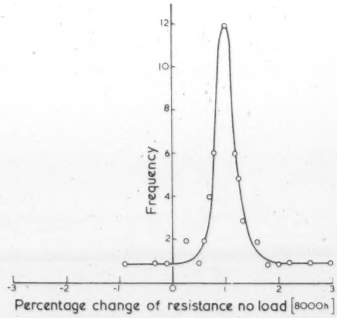


Fig. 3. Frequency diagram showing the percentage resistance changes in resistors after 8000 h at no load.

The annealing or stabilizing treatment appears to bring all the resistors to a common level, and gives a rational basis to resistor production. The on-load stability tests show a slight overall decrease of resistance due to a long-term ordering effect. The increase of resistance at no load is probably due to the effect of water vapour or other contaminating gases in the atmosphere. When the resistors are running hot (on load) these gases do not condense and the resistance value is unaffected. Protection during shelf life can be afforded by the use of special high-temperature resins (e.g. epoxy or silicone resins), of the type normally used in conventional resistor production.*

* We are grateful to the Erie Resistor Co. for carrying out humidity and tropical tests on some films, which showed without doubt that protection is necessary.

Temperature coefficient of resistance

The curve (Fig. 4) shows the t.c.r. within the range 25–110° C plotted against the rate of film deposition (calculated from

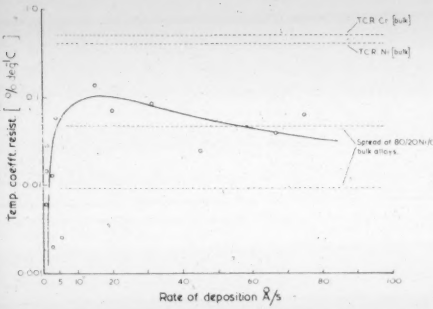


Fig. 4. The temperature coefficient of resistance of nickel–chromium films is shown to be dependent on the rate of deposition. The films studied here were within the range 100–150 Ω/square after annealing.

film thickness and deposition time). These films were within the range 100–150 Ω/square after stabilizing, and were all deposited with a residual gas pressure of 5×10^{-4} mm Hg in the evaporation chamber. The spread of t.c.r. of bulk 80/20 nickel–chromium alloys and the t.c.r. of pure nickel and pure chromium within the same temperature range are also shown in the diagram for comparison.

It is believed that the rate of deposition affects the temperature coefficient of resistance in two ways. Firstly, the films deposited at a slow rate are more highly oxidized by reaction with the residual gas, and secondly, one component of the alloy might preferentially evaporate under the passive evaporation conditions, particularly if the bead becomes unduly large, resulting in a film composition different from the bulk alloy (Siddall and Probyn 1961). These two effects combine to produce a curve of the type shown in Fig. 4. At low rates of deposition, the film is highly oxidized, and the temperature coefficient of resistance is very low. Under extremely low rates of deposition (for example, when a film is deposited by sublimation from a hot nickel–chromium wire) the temperature coefficient is negative. We believe that this can be explained by a gradual change in the electronic conduction mechanism as the films become highly oxidized. When the film is metallic its conductivity is mainly due to the movement of the valency electrons, and the temperature coefficient is positive because of their interaction with the increasing thermal oscillation of the lattice. When the film is oxidized the conductivity may be due to electron defects in the oxide lattice caused by excess metal atoms (i.e. non-stoichiometric composition) or by trivalent cations in a divalent metal oxide lattice (Cr^{+++} and Ni^{++}, for example). Both these models are well known (Kubaschewski and Hopkins 1953) in the theory of semiconducting oxides. The formation of a mixed oxide and metal film would explain the negative t.c.r. observed, and the transition to a rapidly increasing positive value as the true electronic conductivity becomes effective.

As the rate of deposition is increased the deposit becomes more metallic but its composition remains different from the bulk alloy, thus leading to the observed high values of t.c.r. Previous work by Alderson and Ashworth (1957) has shown that the composition of nickel–chromium films evaporated from wire in tungsten heaters is about 60% nickel, 40% chromium and the peak in the t.c.r. curve noted here probably corresponds to such a composition. At high rates of deposition the t.c.r. decreases to that of the bulk material as the alloy evaporates less preferentially.

5. Conclusions

In this paper we have attempted to show that the practical limitations to the use of thin metal films for resistance elements are now understood and that it is possible to make stable resistors by vacuum deposition techniques. It is shown that nickel–chromium films up to some 400 Ω/square on Pyrex glass can be stabilized by annealing, and that the unprotected films remain stable on load, but when off load their resistance value increases by an average of 1% under normal atmospheric conditions.

The low value of temperature coefficient of resistance is an attractive feature of the vacuum-deposited nickel–chromium resistor, and at very low rates of deposition the value approaches zero. The dependence of the temperature coefficient of resistance on the degree of oxidation of the film and fractional distillation of the alloy components can be controlled by the evaporation rate and turned to practical account.

The most important practical requirements of resistor production by vacuum deposition are given below.

(i) A stable glass such as Pyrex should be used for the substrate. This does not rule out the possibility of using glazed or specially prepared ceramics.
(ii) The substrate must be held at a temperature above 200° C and preferably as high as 300° C during deposition, in order to relieve internal stress in the film.
(iii) The oxidation of the film must be roughly controlled by adjustment of the residual gas pressure and the rate of deposition in the vacuum chamber.
(iv) The deposited film must be annealed to stabilize it. This stabilization can be carried out in atmosphere at a temperature in the range 250–350° C.
(v) The deposited film must be protected or encapsulated to improve its stability under widely varying atmospheric conditions.
(vi) Great care is needed in the choice of a suitable method of making contact. Many of the faults arising in thin film and in conventional resistor production can be directly attributed to contact failure.

Acknowledgments

The authors wish to thank Mr. F. D. Edwards, Managing Director, Edwards High Vacuum Ltd., for permission to publish this work and Mr. L. Holland, Head of the Vacuum Deposition Research Division, for his help and encouragement.

References

Alderson, R. H., and Ashworth, F., 1957, *Brit. J. Appl. Phys.*, **8**, 205.

Bassett, G. A., and Pashley, D. W., 1961, *European Regional Conference on Electron Microscopy, Delft* 1960 (Delft: Netherlands Union of Electronmicroscopists).

Fuchs, K., 1938, *Proc. Cambridge Phil. Soc.*, **34**, 100.

Holland, L., and Siddall, G., 1953, *Vacuum*, **3**, 375.

Itterbeeck, A., de Greve, L., and Heremans, F., 1951, *Appl. Sci. Res. B.*, **2**, 320.

KUBASCHEWSKI, O., and HOPKINS, B. E., 1953, *Oxidation of Metals and Alloys* (London: Butterworths), p. 26.
LOVELL, A. C. B., and APPLEYARD, E. T. S., 1937, *Proc. Roy. Soc. A*, **158**, 718.
PLANCK, W., 1914, *Phys. Z.*, **15**, 563.
PRESTON, J. S., GILLHAM, E. J., and WILLIAMS, B. E., 1955, *Phil. Mag.*, **46**, 1051.
SIDDALL, G., and PROBYN, B. A., 1961, *Amer. Vac. Soc. Proc.*, in the press.
THOMSON, J. J., 1901, *Proc. Cambridge Phil. Soc.*, **11**, 120.
TOLANSKY, S., 1948, *Multiple Beam Interferometry of Surfaces and Films* (Oxford: University Press).
VAND, V., 1943, *Proc. Phys. Soc.*, **55**, 222.

Notes and comments

1962 Annual Exhibition of The Institute of Physics and The Physical Society

The 1962 exhibition will be held in the Old and New Halls of the Royal Horticultural Society, Vincent Square, Westminster, London, S.W.1. It will be open at the following times and admission will be by ticket only.

Monday, 15th January	10.30 a.m.–7 p.m. (Members, Students, Subscribers and Group Subscribers of the Institute and Society, and Press only up to 2 p.m.)
Tuesday, 16th January	10 a.m.–9 p.m.
Wednesday, 17th January	10 a.m.–7 p.m.
Thursday, 18th January	10 a.m.–7 p.m.
Friday, 19th January	10 a.m.–1 p.m.

Applications for tickets should be sent to the Secretary of The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, enclosing a stamped, addressed envelope to take the tickets, which measure 4½ in. by 3 in.

The Exhibition *Handbook*, which contains descriptions of exhibits and of some new devices for which space could not be allocated, will be published in mid-December. It is more than a catalogue and serves as a reference book throughout the year. The price is 6*s*. per copy (2*s*. 6*d*. to Members, Students and Subscribers), plus 2*s*. postage.

During the exhibition the following lectures will be given:

Tuesday, 16th January, at 5.45 p.m.

A physicist looks at music, by Dr. C. A. Taylor (Manchester College of Science and Technology).

Thursday, 18th January, at 5.45 p.m.

The uses of elasticity in instrument design, by Professor R. V. Jones (University of Aberdeen).

A leaflet giving abstracts of the above lectures and further information about the exhibition may be obtained from The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1.

Measurement of thermal conductivity of semiconductors at high temperatures

by A. D. STUCKES, M.A., A.Inst.P., Research Department, Associated Electrical Industries (Manchester) Limited, Trafford Park, Manchester 17

MS. received 28th July 1961; paper presented at the Conference on Thermoelectricity held by The Institute of Physics and The Physical Society, at Durham, 10–12th July 1961

Abstract

A brief description is given of absolute, comparative and Ångström methods of measuring thermal conductivity at room temperature and above. Results from various sources on Si, Ge, InSb, InAs and some tellurides are summarized and discussed.

Introduction

The choice of a semiconductor suitable for thermoelectric power generation involves the determination of a number of physical parameters up to temperatures that may be as high as 1000° C. One of these parameters is the thermal conductivity; its evaluation at elevated temperatures is, at present, the most difficult experimental measurement involved.

If a constant heat flow Q is maintained through a sample of cross-sectional area A, and thermal conductivity κ then

$$Q = \kappa A \Delta T / \Delta x$$

where $\Delta T/\Delta x$ is the temperature gradient produced.

In practice Q is difficult to determine accurately owing to extraneous heat losses which increase rapidly with increasing temperature. Also only small samples of experimental semiconductors are generally available and thus great care is needed to make a precise estimate of the temperature gradient.

Present techniques may be divided into two categories: static and dynamic, depending on whether the temperature distribution in the specimen is time-dependent. Static methods may be further subdivided into absolute and comparative. A general description of four methods is given and the results obtained for a number of well-known semiconductors are quoted and compared.

2. Static methods

2.1. *Absolute axial flow*

Absolute methods involve direct measurement of the rate at which heat enters or leaves the specimen. Generally the sample is heated electrically and the power input measured.

In the most common method, which is shown schematically in Fig. 1, the small electric heater produces a temperature gradient axially along the sample. At high temperatures particularly, spurious heat losses from the heater and sample must be avoided to ensure that all the measured heat input passes through the sample. In vacuum these losses will be primarily radiative and a shield must be placed around the heater and the sample, with the temperature gradient in the shield matched to that along the sample.

Thermocouples placed along the sample enable the temperature gradient to be measured. Thermocouple wires and heater leads must be kept as fine as possible to minimize conduction losses along them.

This method has been used very successfully below room temperature. Above room temperature it has been used by Abeles (1959), Busch and Schneider (1954), Deviatkova (1957), Kettel (1959) and Weiss (1958). Experience has

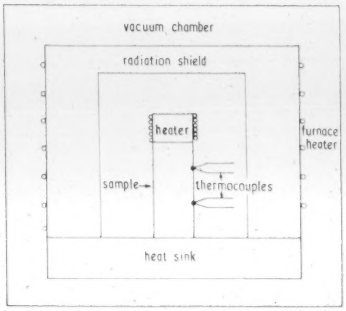


Fig. 1. Absolute axial flow measurement of thermal conductivity.

shown that the prevention of heat loss is extremely difficult, and in some cases results have been in error at the higher temperatures for this reason.

2.2. *Absolute radial flow*

The heat-loss errors associated with the axial flow method can be eliminated by use of a radial heat flow technique which is illustrated in Fig. 2.

Heat is generated along the axis of a cylinder and two thermocouples are placed in the sample at different radii. The thermal conductivity is then computed from the formula

$$\kappa = \frac{Q \ln (r_2/r_1)}{2\pi L \Delta T}$$

where r_1 and r_2 are the radial positions of the inner and outer thermocouples respectively. ΔT is the measured temperature difference for heat energy input Q with a sample length L.

Slack and Glassbrenner (1960) have recently used this method on germanium up to 1020° K. They find theoretically that an error of less than 0·5% occurs from end effects if the length-to-diameter ratio is greater than 4 to 1. Their sample was 1·27 cm in diameter and 6·1 cm long, dimensions very much smaller than those usually advocated for this method. The main problem as the sample becomes smaller is the measurement of the exact radial position of each thermocouple which itself has a finite size. Any error in this

measurement is a constant fraction, independent of temperature, and can be found by measuring the thermal conductivity of the sample by another method at one temperature only and comparing the results.

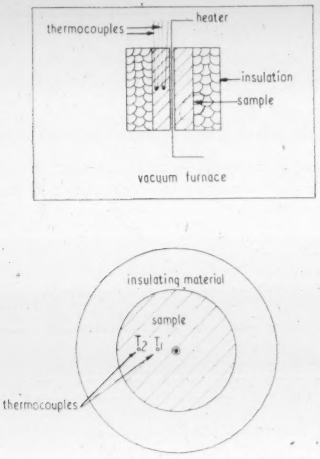


Fig. 2. Schematic diagram of radial flow measurement of thermal conductivity.

2.3. *Comparative method*

In the series comparative method shown in Fig. 3 and described by Stuckes and Chasmar (1956) the sample to be measured is sandwiched between two cylinders of known thermal conductivity. A guard ring reduces lateral heat losses. The temperature gradient in the semiconductor and in both standards is measured and hence the heat flowing into and out of the sample is known. Thus any heat loss is at once obvious and can be taken into account—this is a major advantage in high-temperature measurements.

To minimize lateral heat losses, the temperature drop across the standards and the semiconductor should be kept small and approximately equal—preferably the thermal conductivity of the standards and the semiconductor should be of the same order.

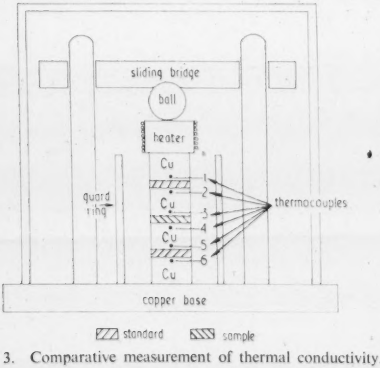


Fig. 3. Comparative measurement of thermal conductivity.

Two thermocouples could be placed in each standard and in the semiconductor. However, the introduction of a thermocouple disturbs the pattern of isothermals and the finite size of the thermocouple makes the exact determination of the thermocouple separation difficult. It would be necessary to use long samples in order to reduce these effects. Also, since semiconductors are normally brittle, mounting thermocouples internally to give a correct indication of body temperature and not surface temperature can be a problem.

These difficulties can be overcome by placing good thermally conducting blocks of silver or copper on either side of the standards and the semiconductor, and placing the thermocouples in these blocks as shown in Fig. 3. This helps to provide a uniform heat flow and, at the same time, the position of the thermocouple is much less critical since the temperature drop within the block is small compared with that in the standards or the semiconductor. This technique, however, introduces the problem of contacts, for it is essential that there is negligible temperature drop at each contact. This has to be solved for each particular semiconductor and usually involves a thin layer of suitable solder. Electrical conductivity measurements can be used as a check on the quality of the thermal contact.

In the experience of the author the contact problem has so far set a temperature limit on this method in the region of 450° C. It is worth remembering that this contact problem is one which also has to be solved in the practical thermoelectric generator.

If a suitable contact cannot be made, longer samples must be used with thermocouples in the semiconductor itself. The guard ring temperature distribution must then be matched over a much greater temperature range.

Stuckes (1957, 1960), Bowers *et al.* (1959), Hust (1961) have measured by the comparative method Ge, Si, InSb, InAs and several tellurides.

3. Dynamic methods

All static measurements are made when the system has come into temperature equilibrium. This may take a period of several hours and is unsuitable where the properties of the specimen change slowly on heating. The techniques of dynamic measurement involve the specimen temperature as a function of time as well as position and can generally be carried out in a few moments. However, in general the methods measure thermal diffusivity, that is the ratio of thermal conductivity and heat capacity per unit volume. Consequently, separate determinations of the specific heat and density are necessary to obtain thermal conductivities. There is a large number of heat capacity values available in the literature and since it does not usually change rapidly with temperature, values extrapolated to the required temperature can often be obtained.

3.1. *Ångström method*

This method, devised originally by Ångström (1863), involves subjecting one end of a bar specimen to a periodic temperature variation. When such a thermal wave is propagated in a medium, the amplitude and phase at any point depend on the properties of the medium and frequency of the wave. Comparison of the amplitude and phase at two points enables the thermal diffusivity to be measured.

It can be shown (Carslaw and Jaeger 1959) that if the distance between two points in a semi-infinite bar is L, τ is the periodic time, β the phase difference and q the ratio of amplitudes, then the thermal diffusivity α is given by

$$\alpha = \frac{\pi L^2}{\tau \beta \ln q}.$$

This assumes (i) that the bar is of such small cross section that the temperature across it does not vary, (ii) that it is losing heat by radiation only to a medium of constant temperature, which is also the initial temperature of the rod, and (iii) that the physical properties of the specimen should be invariant within the range of temperature variations along it.

These assumptions can be achieved fairly closely in practice so that thermal diffusivity can be measured at high temperatures without error due to heat losses.

Sidles and Danielson (1954) used this method to measure metals up to 500° C using very long samples of small diameter. Abeles *et al.* (1960) have recently adapted it to measure germanium up to 800° C. By using a higher frequency they reduced radiation effects and enabled shorter samples to be used. This also reduced the measuring time, which is important since the ambient temperature should not vary during the time required to take a measurement by more than a fraction of the smallest amplitude of temperature measured. Since this is a fraction of a degree, this is a stringent requirement.

The technique of Abeles is illustrated in Fig. 4. Three thermocouples were attached to the sample which was 2 in.

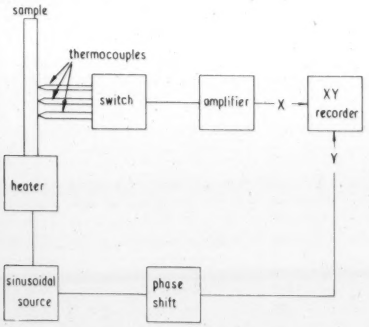


Fig. 4. Measurement of thermal diffusivity by Ångström method.

long by 0·09 in^2 and connected to a switch so that the output from any one thermocouple could be fed through an amplifier. This selected and amplified only the sinusoidal component which was fed to the X plate of an X–Y recorder. A reference voltage from the sinusoidal source supplying the heater was fed through a calibrated phase shift device to the Y input of the recorder. The phase was adjusted until a straight-line Lissajous figure was obtained and thus both the difference in phase and amplitude between adjacent thermocouples could be measured.

The use of a suitable filter eliminated slow temperature variations of the signal, and data could be taken while the furnace temperature was drifting.

Difficulty was encountered in attaching thermocouples to the semiconductor so that they followed the temperature without lag. It was also found that most metallic thermocouples formed eutectics with the semiconductor, well below its melting point. These difficulties were overcome by utilizing the thermal e.m.f. of the semiconductor itself for measuring the relative temperature differences along the sample. Three probes of tungsten were welded along the sample and the e.m.f. between pairs of probes was amplified in place of the thermocouple voltage. Thermal lag was eliminated since electrical contacts only were involved.

In addition to the results for Ge, high-temperature measurements by this technique have been published by Kanai and Nii (1959) for PbTe and InSb.

It seems likely that this method will become increasingly important as measurements are extended to higher temperatures. Even if there is insufficient specific heat data to obtain thermal conductivity, at least the method can be used to find appreciable changes in thermal conductivity since the specific heat changes very little at high temperatures.

4. Results

Although there have been few published measurements of the thermal conductivities of semiconductors above room temperature, comparison of results on the same material indicate the relative reliability of the various methods of measurement. Ideally, lattice thermal conductivities should be compared since the electronic contribution will vary with electrical conductivity and increase with temperature. However, so long as the semiconductors are reasonably pure and the measurements are over a temperature range in which the conduction is extrinsic, the electronic contribution is unlikely to be large.

The measurements discussed are summarized in the Table, which indicates the material, method of measurement and range of temperature involved.

Material	*Method*	*Temperature range, °K*	*Author*
Ge	Absolute axial flow	300–1080	Abeles (1959)
	Absolute axial flow	500–1000	Kettel (1959)
	Absolute radial flow	300–1020	Slack & Glassbrenner (1960)
Si	Comparative	320–700	Stuckes (1960)
	Ångström	300–1070	Abeles *et al.* (1960)
	Comparative	320–570	Stuckes (1960)
	Comparative	300–700	Hust (1961)
InSb	Absolute axial flow	77–770	Busch & Schneider (1954)
	Absolute axial flow	250–700	Weiss (1958)
	Comparative	300–770	Bowers *et al.* (1959)
	Comparative	320–700	Stuckes (1957)
	Ångström	225–450	Kanai & Nii (1959)
InAs	Comparative	300–770	Bowers *et al.* (1959)
	Comparative	320–700	Stuckes (1960)
GaAs	Comparative	320–650	Stuckes (unpublished)
PbTe	Absolute axial flow	100–450	Deviatkova (1957)
	Ångström	110–650	Kanai & Nii (1959)
	Comparative	320–600	Stuckes (unpublished)
CdTe	Comparative	320–600	Chasmar *et al.* (1960)
In_2Te_3	Comparative	320–700	Stuckes (unpublished)

Results for germanium and silicon are given in Fig. 5. Only two measurements on silicon have been reported

(Stuckes 1960, Hust 1961) both using a comparative technique, and reasonable agreement was obtained. Several measurements on germanium have been published. There

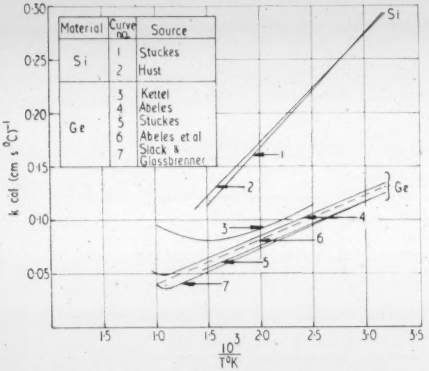


Fig. 5. Thermal conductivity of germanium and silicon.

is remarkably good agreement between the radial flow absolute measurement of Slack and Glassbrenner (1960) and the comparative results (Stuckes 1960). The dotted line, which is very near the above results, indicates the order of magnitude of thermal conductivity obtained from the thermal diffusivity measurements of Abeles (1960). Accurate comparison is not possible owing to lack of specific heat data above room temperature. The room temperature value of specific heat (Flubacher *et al.* 1959) has been used to fix the low temperature end; at 1000° K, which is well above the Debye temperature, an atomic heat at constant volume of 5·96 cal/g atom has been assumed. The results of Abeles (1959) and Kettel (1959) from absolute axial flow methods are somewhat higher. In particular the results of Kettel are of a different form, showing a gradual increase in thermal conductivity from a comparatively low temperature, which cannot be explained in terms of an electronic contribution. It is possible that radiation losses from the heater and sample were not eliminated, an error which would increase with increasing temperature and lead to an apparently high value of thermal conductivity.

Similar, apparently spuriously high, results for indium antimonide have been obtained by Busch and Schneider (1954) and by Weiss (1958) using an absolute method as indicated in Fig. 6. Results from the comparative method and the Ångström method are in good agreement. Results quoted for indium arsenide and gallium arsenide were obtained by comparative methods.

The thermal conductivity of some tellurides is given in Fig. 7. The results for lead telluride obtained by Deviatkova (1957) with an absolute method again show a marked divergence from those obtained by comparative and Ångström methods. The divergence at high temperatures between the latter two results is probably due to the electronic contribution.

Some new results on cadmium telluride and indium telluride have also been included.

5. Conclusions

Examination of the results for Ge and InSb which have been measured by many techniques, indicate that the comparative, the radial flow and the Ångström methods are in good agreement. Where results using other techniques are not available (Si, InAs) measurements by different workers using the comparative method are consistent with each other.

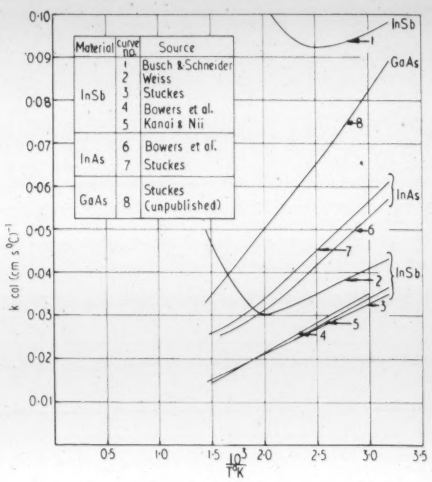


Fig. 6. Thermal conductivity of some group III–V compounds.

This method would thus seem reliable and most suitable where only small specimens can be obtained. Using larger specimens, the absolute radial flow method is probably to be preferred since it enables measurements to be taken to higher

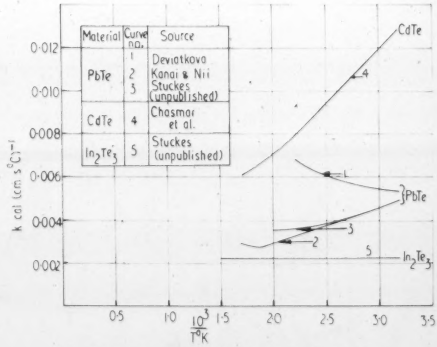


Fig. 7. Thermal conductivity of some tellurides.

temperatures without an associated thermal contact problem.

The Ångström method has the advantage of speed and, if specific heat data are available, is suitable for measurements to very high temperatures.

Acknowledgments

The author wishes to thank Mr. R. P. Chasmar for his continued interest in this subject and to thank Sir Willis

Jackson, Director of Research, Associated Electrical Industries (Manchester) Ltd., for permission to publish this paper.

References

ABELES, B., 1959, *J. Phys. Chem. Solids*, **8**, 340.
ABELES, B., CODY, G. D., and BEERS, D. S., 1960, *J. Appl. Phys.*, **31**, 1585.
ÅNGSTRÖM, A. J., 1863, *Phil. Mag.*, **25**, 130; **26**, 161.
BOWERS, R., URE, R. W., BAUERLE, J. E., and CORNISH, A. J., 1959, *J. Appl. Phys.*, **30**, 930.
BUSCH, G., and SCHNEIDER, M., 1954, *Physica*, **20**, 1084.
CARSLAW, H. S., and JAEGER, J. C., 1959, *Conduction of Heat in Solids*, 2nd edn. (Oxford. Clarendon Press), p. 136.
CHASMAR, R. P., DURHAM, E. D., and STUCKES, A. D., 1961, *Proceedings of the International Conference on Physics of Semiconductors, Prague* 1960 (Prague: Publishing House of Czechoslovak Academy of Sciences), p. 1018.
DEVIATKOVA, E. D., 1957, *Soviet Phys. Tech. Phys.*, **2**, 414.
FLUBACHER, P., LEADBETTER, A. L., and MORRISON, J. A., 1959, *Phil. Mag.*, **4**, 273.
HUST, J. G., 1961, *Tech. Rep.* No. 2, Contract Nonr. 2964(01). Office of Naval Research.
KANAI, Y., and NII, R., 1959, *J. Phys. Chem. Solids*, **8**, 338.
KETTEL, F., 1959, *J. Phys. Chem. Solids*, **10**, 52.
SIDLES, P. H., and DANIELSON, G. C., 1954, *J. Appl. Phys.*, **31**, 1585.
SLACK, G. A., and GLASSBRENNER, C., 1960, *Phys. Rev.*, **120**, 782.
STUCKES, A. D., 1957, *Phys. Rev.*, **107**, 427.
—— 1960, *Phil. Mag.*, **5**, 84.
STUCKES, A. D., and CHASMAR, R. P., 1956, *Report of the Meeting on Semiconductors* (London: The Physical Society), p. 119.
WEISS, H., 1958, *Semiconductors and Phosphors* (Braunschweig: Vieweg), p. 497.

Notes and comments

Society of Environmental Engineers

The Society of Environmental Engineers, now in its third year, provides a forum for discussion of problems and the latest techniques in environmental testing, and organizes visits, symposia and exhibitions in this field.

Meetings are held in the Mechanical Engineering Department, Imperial College, Exhibition Road, London, S.W.7, beginning at 6 p.m. Branch meetings in different parts of the country have also been arranged.

The annual proceedings of the Society for 1959–60 and 1960–61 are now available (price 25*s*. each). Further details and membership forms can be obtained from the Hon. Secretary, Suite 7, 167 Victoria Street, London, S.W.1.

London International Engineering Exhibition, 1963

The London International Engineering Exhibition, formerly the Engineering, Marine, Welding and Nuclear Energy Exhibition, will be held at Olympia and Earls Court from 23 April to 2 May, 1963. It will be organized by F. W. Bridges and Sons Ltd.

The heavier exhibits, including a welding section, will on the whole be displayed at Olympia and electrical engineering and lighter products at Earls Court.

Further information may be obtained from James Brewster, Public Relations Officer, Press Office, Commonwealth House, New Oxford Street, London, W.C.1.

Colour Group (Great Britain)

The Colour Group (Great Britain) arranges about eight meetings a year for those interested in all aspects of colour and related aspects of vision. Visits, conferences and exhibitions are also held from time to time. All enquiries should be addressed to the Hon. Secretary, Mr. F. J. B. Wall, Ilford Ltd., Brentwood, Essex.

This was originally one of the specialist groups of the Physical Society and it is hoped it will have a successful future.

Second International Conference of Cybernetic Medicine, 1962

The Second International Conference of Cybernetic Medicine will be held in Amsterdam, The Netherlands, from 16–19 April, 1962.

The main sections of the Congress will be: Symposium on cybernetics of the nervous system; Symposium on cybernetics of endocrine glands; Free communications.

Further information can be obtained from Dr. J. P. Schadé, Secretary General, Netherlands Central Institute for Brain Research, Mauritskade 59b, Amsterdam, The Netherlands.

Rapid method for determining v.h.f. dielectric parameters for liquids and solutions using standing wave procedures

by M. G. CORFIELD, B.Sc., J. HORZELSKI, and A. H. PRICE, Ph.D., The Edward Davies Chemical Laboratory, University College of Wales, Aberystwyth

MS. received 3rd August 1961

Abstract

A rapid method of measuring the dielectric parameters of liquids and solutions in the frequency range 250 *to* 900 *Mc/s is described. The method adopted is a variation of the one due to Roberts and von Hippel, but involves far less mathematical labour. Permittivities may be measured to an accuracy of* $\pm 2\%$ *and loss tangents to* $\pm 1 \times 10^{-3}$ *or* $\pm 5\%$ *(whichever is the greater).*

Introduction

At frequencies above about 250 Mc/s the dielectric parameters of materials are measured using distributed circuits, and the results often calculated by the method of Roberts and von Hippel (1946, von Hippel 1954, p. 63). This involves measuring the standing wave pattern set up in the coaxial line or wave-guide when the imput signal is reflected at a short circuit termination placed immediately behind the dielectric. The method and experimental procedure has been described by Williams (1959). One disadvantage with this general method is the large amount of mathematical labour involved in calculating the dielectric parameters from the standing wave pattern. Thus, the equation

$$\frac{\tanh \gamma_2 d}{\gamma_2 d} = \left\{ \frac{s - j \tan (2\pi x_0/\lambda_1)}{j\beta_1 d(1 - js \tan (2\pi x_0/\lambda_1))} \right\} \qquad (1)$$

(where γ_2 is the propagation constant of a dielectric sample of depth d, s the inverse voltage standing wave ratio, x_0 the distance of the first node of the standing wave pattern in air from the dielectric surface, λ_1 the free space wavelength of the radiation, β_1 the phase constant of the incident radiation, and $j = \sqrt{-1}$) has to be solved for γ_2. Charts are available for this purpose (von Hippel 1954, p. 86), and the sample permittivity ε' and loss tangent $\tan \delta$ may then be calculated.

Dakin and Works (1947) have shown that for medium and low loss materials Eqn (1) may be written as

$$\frac{\tan \beta_2 d}{\beta_2 d} = -\left(\frac{\lambda_1}{2\pi d}\right) \tan \left(\frac{2\pi x_0}{\lambda_1}\right)$$

$$\tan \delta = \left(\frac{\Delta x}{d}\right) \left[\frac{\beta_2 d\{1 + \tan^2 (2\pi x_0/\lambda_1)\}}{\beta_2 d(1 + \tan^2 \beta_2 d) - \tan \beta_2 d} \right] \qquad (2)$$

(after separating the real and imaginary components) where β_2 is the phase constant of the dielectric, and Δx the contribution of the dielectric to the width (at twice the minimum power) of the standing wave in air. If, however, the depth of the sample is adjusted to an integral number of half wavelengths, or an odd integral number of quarter wavelengths of the radiation *in the sample*, the mathematical labour becomes trivial (von Hippel 1954, p. 119). Thus, for an integral number of half wavelengths in the sample the permittivity ε' and loss tangent $\tan \delta$ are given by

$$\varepsilon' = \left(\frac{m\lambda_1}{2d}\right)^2 \qquad (3a)$$

and

$$\tan \delta = \frac{\Delta x}{d} \qquad (3b)$$

where m is the number of half wavelengths in the dielectric of depth d. Similarly, for an odd number of quarter wavelengths *in the sample*

$$\varepsilon' = \left(\frac{n\lambda_1}{4d}\right)^2 \qquad (4a)$$

and

$$\tan \delta = \frac{\Delta x}{\varepsilon' d} \qquad (4b)$$

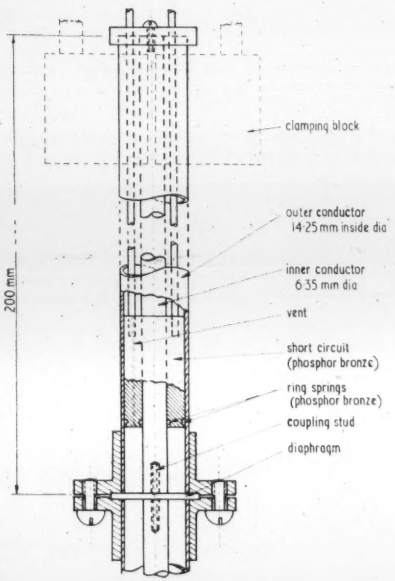
where n is the number of quarter wavelengths in the sample.

Apparatus

A General Radio (GR) oscillator (type 1209B) of frequency range 250–920 Mc/s injects a singal into a horizontally placed GR874-LBA slotted line of 50 ohms characteristic impedance. The measuring cell (see Figure) is attached vertically to the slotted line by a right angle joint (GR type 874EL) and GR874 coaxial connectors. The standing wave pattern in the slotted line is measured using a probe as described by Williams (1959).

The measuring cell is constructed from a 30 cm length of coaxial line (50 ohms characteristic impedance) cut about 10 cm from one end. Flanges are built on the outer conductor on either side of the cut, and these hold a thin (about 0·3 mm) polythene (or other non-polar, inert material) diaphragm in position. The inner conductor is reassembled by a coupling stud passing through a small hole in the diaphragm. This arrangement forms a liquid tight seal. A phosphor bronze short circuit moves inside the coaxial line

making electrical contact with the inner and outer conductors by two phosphor bronze ring springs. A small vent, drilled in the short circuit, allows displacement of the liquid on moving the short circuit. Two rigid rods connect the short circuit to a coarse and fine adjusting mechanism, whose movement is monitored by a cathatometer focused on a reference mark.



Measuring cell: detail of short circuit and diaphragm.

Experimental procedure at a given frequency, and calculations of the results

Half wavelength measurements

With the short circuit removed the cell is filled with liquid. The former is replaced and moved until in contact with the diaphragm. The position of the standing wave voltage minimum in the slotted line and its width Δx_0 at twice minimum power are measured using the probe. This width is a measure of the attenuation due to the system between the probe and the bottom of the liquid column. The short circuit is then moved through the liquid until a voltage minimum reappears at the probe, and the distance moved d equals a half wavelength for the radiation in the liquid. This distance, substituted in Eqn (3*a*), yields the permittivity of the dielectric. The width Δx_1 of the voltage minimum is also measured, and the difference between this reading and the one obtained with the short circuit at the diaphragm gives the apparent contribution due to the loss in the liquid (i.e. $\Delta x = \Delta x_1 - \Delta x_0$). By substitution in Eqn (3*b*) the apparent loss tangent of the liquid is determined, and when corrected for any wall loss in measuring cell gives the liquid loss tangent. The procedure may then be repeated for any further half wavelengths in the liquid.

Quarter wavelength measurement

Here the probe is set at an integral number of quarter wavelengths from the diaphragm and the position of the short circuit adjusted until a voltage minimum appears at the probe when the cell is filled with liquid. The distance d between the diaphragm and short circuit (measured in the same way as for the half wavelength determination) is substituted in Eqn (4*a*), thus giving the permittivity of the sample. The width of the voltage minimum is also measured, and when corrected for any losses in the slotted line and the diaphragm, gives the value of Δx to be substituted in Eqn (4*b*). The resulting loss tangent is finally corrected for any wall loss in the cell.

The calculation of the dielectric parameters may be illustrated with the results for water. Due to the permittivity ε' term in the denominator of the expression for tan δ in the quarter wave Eqn (4*b*), the measured Δx is greater for quarter wave measurements than it is for half wavelength measurements. For water ($\varepsilon' = 80{\cdot}4$) the Δx values are so large at a quarter wavelength that accurate measurements are only possible at half wavelengths in the liquid. The experimental determinations and calculation of the results for water at 900 Mc/s and 20° C are shown in Table 1.

Table. 1. *Calculation of the results for water at* 20° C *and* 900 *Mc/s*

Wavelength incident radiation in air $= \lambda_1 = 33{\cdot}48$ cm

m	*Position of short circuit*	d (*cm*)	$\Delta x(m)$ (*cm*)	Δx (*cm*)	$\varepsilon' = \left(\frac{m\lambda_1}{2d}\right)^2$	tan $\delta \times 10^3$
0	35·04	0	0·132	0		
1	36·90	1·86	0·226	0·094	81·0	50·5
2	38·77	3·73	0·329	0·197	80·6	52·8
3	40·65	5·61	0·429	0·297	80·1	52·9
4	42·52	7·48	0·519	0·387	80·1	51·7
5	44·39	9·35	0·603	0·471	80·1	50·4
6	46·26	11·22	0·691	0·559	80·1	49·8

m = number of half wavelengths in sample; d = distance moved by short circuit from $m = 0$; $\Delta x(m)$ = width voltage minimum; $\Delta x = \Delta x(m = m) - \Delta x(m = 0)$; tan $\delta = \Delta x/d$. Mean $\varepsilon' = 80{\cdot}3 \pm 0{\cdot}7$; mean tan $\delta = (51{\cdot}4 \pm 1{\cdot}6) \times 10^{-3}$.

This measured loss tangent, however, includes a contribution due to a wall loss of $1{\cdot}0 \times 10^{-3}$ in the measuring cell. This must be subtracted from the measured loss in order to obtain the true loss tangent of the water. Hence, tan δ (water) $= (50{\cdot}4 \pm 1{\cdot}6)10^{-3}$. These values agree well with the literature ones (see below).

Experimental results

The experimental results for a variety of systems (together with the corresponding literature values) are given in Table 2. The systems were chosen to illustrate cases of high and low permittivities, and also medium and low loss tangents. All materials were purified before use and their physical constants agreed with the literature values.

The literature values were calculated from the data tabulated by Buckley and Maryott (1958).

It is seen from the table that good agreement is obtained between the experimentally determined values and those quoted in the literature.

Table 2. *Experimental determinations at* 20° C

System	Quarter or half wavelength in sample	Frequency Mc/s	Experimental values ε′	Experimental values tan δ × 10³	Literature values ε′	Literature values tan δ × 10³
Water	Half	900	80·3 ± 0·7	50 ± 2	80·4	49·5
		350	81·0 ± 0·2	19 ± 1	80·4	19·4
Chlorobenzene	Quarter	900	5·68 ± 0·02	39 ± 2	5·69	36·5
		350	5·62	14	5·69	14·2
	Half	900	5·74 ± 0·05	39 ± 2	5·69	36·5
		400	5·70 ± 0·05	17 ± 2	5·69	16·3
Benzene	Quarter	900	2·28	1·1	2·28	0
		350	2·26	1·6	2·28	0
	Half	900	2·28	0	2·28	0
		640	2·28	0	2·28	0
Benzophenone in benzene (10·58 g/l.)	Quarter	900	2·37	2·5		2·8
		350	2·37	1·6		1·1

Conclusions

At present the apparatus is capable of determining permittivities to an accuracy of about ±2%, and loss tangents to at least $\pm 1 \times 10^{-3}$ or ±5% (whichever is the greater).

The method has a number of advantages over the general procedure of Roberts and von Hippel. The time taken for the experimental determination and calculation is very much reduced. Thus, measurement and calculation at three frequencies using the general method often take up to six hours, whereas the same information is obtained in about two hours by the method described in this paper. Another difficulty with the general method is that the liquid meniscus introduces some uncertainty into the determination of the correct liquid depth, and hence increases the error in the measured dielectric parameters. This error increases with increasing permittivity of the liquid. In the apparatus described the dielectric is supported on a diaphragm and the short circuit immersed in the liquid, and these meniscus effects are thus eliminated.

The apparatus may be used over any frequency range where standing wave methods are employed for dielectric determinations.

Acknowledgments

The authors wish to thank Dr. Mansel Davies and Mr. R. Bennett (of the Royal Technical College, Salford) for many helpful discussions. One of us (M. G. C.) wishes to thank the Department of Scientific and Industrial Research for a Research Studentship.

References

BUCKLEY, F., and MARYOTT, A. A., 1958, *Tables of Dielectric Dispersion Data for Pure Liquids and Dilute Solutions* (Washington: National Bureau of Standards, Circular 589).

DAKIN, T. W., and WORKS, C. N., 1947, *J. Appl. Phys.*, **18**, 789.

VON HIPPEL, A., 1954, *Dielectric Materials and Applications* (New York: John Wiley).

ROBERTS, S., and VON HIPPEL, A., 1946, *J. Appl. Phys.*, **17**, 610.

WILLIAMS, G., 1959, *J. Phys. Chem.*, **63**, 534.

On the transfer relation of a manometer for liquid level measurements

by S. DUTTA ROY, M.Sc.(Tech.), River Research Institute, West Bengal, India

MS. received 5th June 1961

Abstract

The relation between the motion of a mercury surface in a manometer for liquid level measurements, and that of the liquid to which it is connected, has been deduced and methods have been worked out for making this relation linear.

1. Introduction

THE principle of a manometer is well known in connection with the measurement of gas pressures. Recently Ray, Dutta Roy and Daskaviraj (1961) devised a simple method for its application in measuring a varying water level. The apparatus is shown in Fig. 1. The

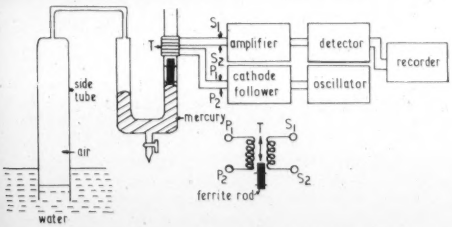


Fig. 1. Water level recorder utilizing the principle of a manometer. (Ray, Dutta Roy and Daskaviraj 1961.)

U-tube is initially filled with mercury up to an appropriate height and the side tube is then immersed in water whose level is to be measured. As the water level rises the air enclosed in the side tube and part of the U-tube is compressed, causing a difference in level of mercury in the two arms of the U-tube. A ferrite rod floats in the arm of the U-tube open to the atmosphere and acts as a core of variable depth of immersion in a transformer T wound on the surface of the U-tube. A constant audio-frequency voltage is injected into the primary, and the induced voltage in the secondary, which varies according to level fluctuations, is amplified, detected and recorded.

Although the overall response of the system depends on the manometer as well as the transformer, a linear relation for the former is highly desirable for the many possible uses of the manometer, with or without the transducing arrangement.

2. The transfer relation of the manometer

For deducing the transfer relation of the manometer, we shall consider the simplified system of Fig. 2. In this, Q and R are two points on the same horizontal plane as the lower face of the side tube and S and T are similar points on the two arms of the U-tube. ρ_1 and ρ_2 denote the densities of liquids 1 and 2 respectively, h_e and h_i are respectively the rises of liquid level outside and inside the

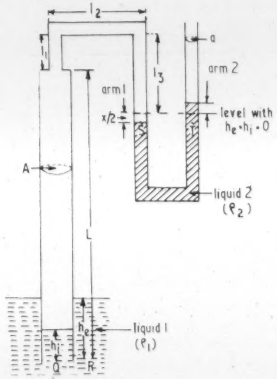


Fig. 2. Showing the symbols adopted for analysis.

side tube, A and a are the areas of cross section of side tube and U-tube respectively (both internal, assumed to be uniform throughout), L is the total length of the side tube and $l = l_1 + l_2 + l_3$ is the length of the U-tube from the upper end of the side tube to the original level (i.e. when $h_e = h_i = 0$) of liquid 2 in arm 1 of the U-tube, $\frac{1}{2}x$ is the depression in arm 1 which is also equal to the rise in arm 2 due to a rise h_e of liquid 1.

By the transfer relation of the manometer, we mean the relation between h_e and x. Let P and P' denote respectively the atmospheric pressure and the pressure of the gas enclosed, in absolute units. P_Q, P_R, P_S and P_T are the pressures at the points Q, R, S and T respectively, in the same units. Let

$$P = H\rho_2 g \tag{1}$$

and

$$\lambda = \rho_1/\rho_2.$$

Then since $P_S = P_T$ and $P_Q = P_R$, we have the two relations

$$P' = P + x\rho_2 g \tag{2}$$

$$P' + h_i\rho_1 g = P + h_e\rho_1 g. \tag{3}$$

From Eqns (2) and (3) we have

$$P + x\rho_2 g + h_i\rho_1 g = P + h_e\rho_1 g$$

or

$$x = (h_e - h_i)\lambda. \tag{4}$$

Eqn (4) expresses the transfer relation of the manometer in terms of an unknown quantity h_i. To a first approximation, we can assume that $h_i \ll h_e$, so that

$$x \simeq \lambda h_e. \tag{5}$$

To derive an accurate expression, h_i must be eliminated from Eqn (4). We can do this by using Boyle's law for gases. We shall assume that the variations in level are slow so that the gas is subjected to isothermal changes only. The adiabatic case will be discussed in § 5.

The initial pressure of the gas enclosed is P and the initial volume is $AL + al$. The final pressure and volume are respectively $P + x\rho_2 g$ and $(L - h_i)A + (l + \frac{1}{2}x)a$. Therefore, by Boyle's law,

$$P(AL + al) = (P + x\rho_2 g)\{(L - h_e + x/\lambda)A + (l + \tfrac{1}{2}x)a\} \tag{6}$$

where h_i has been expressed in terms of h_e and x from Eqn (4). Eqn (6) can be simplified as

$$P\{(A/\lambda + \tfrac{1}{2}a)x - Ah_e\} + x\rho_2 g\{AL + al - Ah_e + (A/\lambda + \tfrac{1}{2}a)x\} = 0. \tag{7}$$

Combining Eqns (7) and (1) and simplifying, we have

$$x^2(A/\lambda + \tfrac{1}{2}a) - xh_eA + x\{AL + al + H(A/\lambda + \tfrac{1}{2}a)\} - h_eAH = 0. \tag{8}$$

This represents a hyperbola and is an accurate expression for the manometer transfer relation. In order to see how x varies with h_e, we write Eqn (8) in the form

$$h_e = \frac{x[(A/\lambda + \tfrac{1}{2}a)x + \{AL + al + H(A/\lambda + \tfrac{1}{2}a)\}]}{A(H + x)}. \tag{9}$$

For values of $x \ll H$, the second term in the numerator far outweighs the first term, so that

$$h_e \simeq \frac{x\{AL + al + H(A/\lambda + \tfrac{1}{2}a)\}}{AH}. \tag{10}$$

Non-linearity in the curve of h_e against x will thus be shown at higher values of x only.

3. Conversion of the non-linear relation to a linear one

The non-linear relation (8) can be converted to a perfectly linear one by using a side tube of non-uniform cross section, i.e. by using a side tube whose cross section at a particular height h_e is a function of h_e, and correspondingly of h_i, say $\phi(h_i)$.

From Eqn (4), we note that if h_e is linearly related to h_i, then a linear relation will be obtained between x and h_e also. Thus if

$$h_e = \mu h_i \tag{11}$$

we shall have

$$x = \lambda(\mu - 1)h_i = \nu_i h_i \tag{12a}$$

$$= \frac{\lambda(\mu - 1)h_e}{\mu} = \nu_e h_e \tag{12b}$$

where the constants ν_i and ν_e are both positive because $\mu > 1$.

With a side tube of non-uniform cross section, the relation (6) is modified to

$$P\left\{\int_0^L \phi(h_i)dh_i + al\right\} = (P + x\rho_2 g)\left\{\int_{h_i}^L \phi(h_i)dh_i + (l + \tfrac{1}{2}x)a\right\}. \tag{13}$$

Substituting for P and x from Eqns (1) and (4), we have

$$H\left\{\int_0^L \phi(h_i)dh_i + al\right\} = \{H + \lambda(h_e - h_i)\}\left[\int_{h_i}^L \phi(h_i)dh_i + a\{l + \tfrac{1}{2}\lambda(h_e - h_i)\}\right]. \tag{14}$$

Let

$$V = \int_0^L \phi(h_i)dh_i$$

represent the total internal volume of the side tube. The left-hand side of Eqn (14), $H(V + al)$ is a constant, equal to β, say. Also to find the form of $\phi(h_i)$ for a linear relation between x and h_e, we substitute in Eqn (14) the desired relation (11). We obtain

$$\{H + \lambda(\mu - 1)h_i\}\left[\int_{h_i}^L \phi(h_i)dh_i + a\{l + \tfrac{1}{2}\lambda(\mu - 1)h_i\}\right] = \beta. \tag{15}$$

From Eqns (15) and (12), we have

$$(H + \nu_i h_i)\left\{\int_{h_i}^L \phi(h_i)dh_i + a(l + \tfrac{1}{2}\nu_i h_i)\right\} = \beta$$

or

$$\int_{h_i}^L \phi(h_i)dh_i = \frac{\beta}{(H + \nu_i h_i)} - (\eta + \delta h_i) \tag{16}$$

where $\eta = al$ and $\delta = \frac{1}{2}a\nu_i$. Let

$$\int \phi(h_i)dh_i = F(h_i) + K_0 \tag{17}$$

where K_0 is a constant. Then from Eqns (16) and (17), we have

$$F(L) - F(h_i) = \frac{\beta}{(H + \nu_i h_i)} - (\eta + \delta h_i). \tag{18}$$

Differentiating both sides of (18), we have

$$F'(h_i) = \phi(h_i) = \frac{\beta\nu_i}{(H + \nu_i h_i)^2} + \delta.$$

For convenience, we write $\phi(h_i)$ in the form

$$\phi(h_i) = \alpha_0 + \frac{\alpha_1}{(\alpha_2 + \alpha_3 h_i)^2} \tag{19}$$

where

$$\alpha_0 = \delta = \tfrac{1}{2}a\lambda(\mu - 1)$$
$$\alpha_1 = \beta\nu_i = H\lambda(\mu - 1)(V + al)$$
$$\alpha_2 = H$$

and

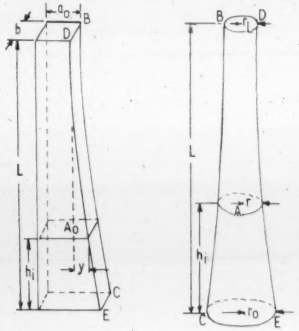
$$\alpha_3 = \nu_i = \lambda(\mu - 1).$$

Eqn (19) gives the nature of non-uniformity required for making the transfer relation perfectly linear.

4. Practical realization

We next consider the question of how physical structures can be made to obey Eqn (19). A look at this equation reveals two facts: (i) the cross section has a constant and a variable term; (ii) the cross section decreases with increasing h_i. Two simple cases will be considered: Case I. Rectangular cross section. Case II. Circular cross section.

Case I. The physical shape is shown in Fig. 3. With the symbols adopted in this figure, we require to know the



(*a*) (*b*)

Fig. 3. Non-uniform side tubes that can be used for linearization of the transfer relation (*a*) case I, rectangular cross section; (*b*) case II, circular cross section.

relation between y and h_i for practical construction of the tube. A typical cross section A at a height h_i has the area

$$\phi(h_i) = b(a_0 + y) = a_0 b + by. \tag{20}$$

Equating the right-hand sides of Eqns (19) and (20), we can write

$$a_0 b = \alpha_0,$$

and

$$by = \frac{\alpha_1}{(\alpha_2 + \alpha_3 h_i)^2}$$

or,

$$y(\alpha_2 + \alpha_3 h_i)^2 = \alpha_1 / b \tag{21}$$

which gives the desired relation between hi and y. DE (or BC) will therefore be a third degree curve.

Case II. The physical shape is shown in Fig. 3(*b*). A typical cross section A has the area

$$\phi(h_i) = \pi r^2 \tag{22}$$

where r is the radius at a height h_i. From Eqns (22) and (19), we have

$$r^2 = \frac{\alpha_0}{\pi} + \frac{\alpha_1}{\pi(\alpha_2 + \alpha_3 h_i)^2}. \tag{23}$$

This gives the equation to the curve DE (or BC), which is now a fourth degree one. The maximum and minimum radii are given by

$$\left.\begin{aligned} r_{max} &= r_0 \left\{\frac{(\alpha_0 + \alpha_1/\alpha_2^2)}{\pi}\right\}^{\frac{1}{2}} \\ r_{min} = r_L &= \left[\left\{\alpha_0 + \frac{\alpha_1}{(\alpha_2 + \alpha_3 L)^2}\right\}\frac{1}{\pi}\right]^{\frac{1}{2}} \end{aligned}\right\}. \tag{24}$$

5. Transfer relation for the adiabatic case

When the level fluctuations are rapid, the analysis of § 2 does not hold, because Boyle's law applies to the isothermal case only. For the adiabatic case, Eqn (6) modifies to

$$H(AL + al)^\gamma = (H + x)\{L - h_e + x/\lambda)A + (l + \tfrac{1}{2}x)a\}^\gamma \tag{25}$$

where γ is the ratio C_p/C_v for the gas enclosed: C_p and C_v are the specific heats under constant pressure and volume respectively. Eqn (25) can be simplified as

$$H(AL + al)^\gamma = (H + x)\{AL + al + x(A/\lambda + \tfrac{1}{2}a) - Ah_e\}^\gamma$$

or,

$$\left(1 + \frac{x}{H}\right)\left\{1 + \frac{x(A/\lambda + \tfrac{1}{2}a)}{AL + al} - \frac{h_e A}{AL + al}\right\}^\gamma = 1. \tag{26}$$

Let

$$\frac{A/\lambda + \tfrac{1}{2}a}{AL + al} = m \quad \text{and} \quad \frac{A}{AL + al} = n.$$

We note that both m and n are less than unity. In fact, $m \ll 1$, and $n < 1$ for practical cases. Eqn (26) now transforms to

$$(1 + x/H)\{1 + (mx - nh_e)\}^\gamma = 1$$

or,

$$\left(1 + \frac{x}{H}\right)\left\{1 + \gamma(mx - nh_e) + \frac{\gamma(\gamma - 1)}{2!}(mx - nh_e)^2 + \ldots\right\} = 1.$$

Neglecting the square and other higher power terms, we have

$$(1 + x/H)\{1 + \gamma(mx - nh_e)\} = 1$$

or,

$$x^2\gamma m - xh_e\gamma n + x(1 + \gamma mH) - h_e\gamma nH = 0. \tag{27}$$

Eqn (27) which gives the transfer relation for the adiabatic case is of the same form as Eqn (8) of § 2 and the same remarks apply here also. Again the coefficients of x and h_e predominate and non-linearity occurs at high values of h_e only.

6. Linearization for the adiabatic case

For the adiabatic case, Eqn (15) is modified to

$$(H + v_i h_i)\left\{\int_{h_i}^{L} \phi(h_i)dh_i + a(l + \tfrac{1}{2}v_i h_i)\right\}^\gamma = K$$

where

$$K = H\left\{\int_0^L \phi(h_i)dh_i + al\right\}^\gamma = \text{constant}.$$

Therefore,

$$\int_{h_i}^{L} \phi(h_i)dh_i = \frac{K}{(H + v_i h_i)^{1/\gamma}} - (\eta + \delta h_i). \tag{28}$$

It can be shown that $\phi(h_i)$ is of the form

$$\phi(h_i) = \omega_0 + \frac{\omega_1}{(\omega_2 + \omega_3 h_i)^{1 + \frac{1}{\gamma}}} \tag{29}$$

where ω_0, ω_1, ω_2, ω_3 are constants. The equations for the physical structures corresponding to the cases I and II of § 4 can be worked out in a similar manner from Eqn (29).

In case I, the resulting equation will be of $(2 + 1/\gamma)$th degree while that for case II will be of degree $3 + 1/\gamma$.

Acknowledgments

This work is published with the kind permission of Shri B. Maitra, Director, River Research Institute, West Bengal, India. The author thanks Shri N. Majumdar for helpful discussions.

Reference

RAY, P. B., DUTTA ROY, S., and DASKAVIRAJ, S. K., 1961, *Irrigation and Power* (India), **18,** 138.

Notes and comments

British Standards

We have recently received the following British Standards: B.S. 526: 1961, Definitions of the calorific value of fuels; B.S. 1193: 1961, Specification for sizes of sensitized materials for recording instruments; B.S. 1928: 1961, Specification for gramophone records and reproducing equipment.

These are available, price 6*s*., from British Standards House, 2 Park Street, London, W.1.

Notes on the British Iron and Steel Association Open Days

The British Iron and Steel Association has compiled a set of notes on the exhibits which were displayed at their open days at Sheffield in June, for the information both of those who were able to attend the open days and to others interested in the Association's work.

Copies of these notes are available, free of charge, from The Information Officer, Development and Information Services, British Iron and Steel Association, 11 Park Lane, London, W.1.

Diamonds in Industry

We have received a copy of the booklet Diamonds in Industry which has been published by the Industrial Diamond Information Bureau. It has been prepared for students and engineers and contains information under the headings of The hardness of diamond, Gem polishing, Grinding wheels, Glass working, Dressing and truing, Honing and lapping, Diamond saws, Diamonds for drilling, Diamonds for lathe tools and wire drawing dies, Non-cutting uses of diamond and New applications.

Copies may be obtained, free of charge, from the Industrial Diamond Information Bureau, 2 Charterhouse Street, London, E.C.1.

Conference on 'The Ionosphere'

The Institute of Physics and The Physical Society announces that it is arranging a conference on 'The Ionosphere' to take place in London from 2nd–6th July, 1962. It is intended as a working conference and no social activities are planned. The subjects to be covered and the opening speakers are: Ionospheric Constitution and Ionizing Radiations by Dr. H. Friedman of the U.S. Naval Research Laboratory, Washington; Geomagnetism and the Ionosphere by Dr. C. O. Hines of the Defence Research Telecommunications Establishment, Ottawa; Irregularities in the Ionosphere by Dr. B. H. Briggs of the Cavendish Laboratory, Cambridge; Mathematics of Wave Propagation through the Ionosphere by Professor H. G. Booker of Cornell University.

Each of these principal sessions will be opened by the named speaker and offers of other contributions are invited. These should be accompanied by an abstract of 100–150 words each and be sent to the Chairman of the Organizing Committee, Mr. J. A. Ratcliffe, at the Radio Research Station, Ditton Park, Slough, Bucks, before 1st March next. The final scripts will be required by 1st June, 1962, and it is hoped that publication of the proceedings of the conference will take place very soon after the close of the conference. The official languages of the conference will be English and French but whenever possible papers and abstracts should be presented in English. The organizers are not proposing at present to make any special arrangements for accommodation for those attending the conference. Those wishing to do so are strongly advised to book hotel accommodation in London very early in the New Year. Foreign visitors would be assisted by the London Hotels Information Service, 88 Brook Street, London, W.1.

Further particulars will be available at the end of March from the Administration Assistant, The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1.

Investigation of a method of growing crystals of GaP and GaAs from the vapour phase

by G. R. ANTELL, B.Sc., A.M.I.E.E., Research Department, Associated Electrical Industries (Manchester) Ltd., Trafford Park, Manchester 17

MS. received 18th July 1961

Abstract

Crystals of GaP and GaAs were grown from the vapour at about 900° C using an iodide process. The crystals were deposited on a localized cold spot and their initial growth could be readily observed. Single crystal samples of GaP had resistivities of the order of 5×10^9 *ohm cm. The concentration of iodine in the GaP crystals was found to be less than 2 parts in* 10^8. *A p-type sample of GaAs was grown at less than 800° C which had a Hall coefficient of 53* cm^3 $coulomb^{-1}$.

1. Introduction

The growth of small crystals of the semiconducting compounds GaP, GaAs, InP and InAs using a technique based on the disproportionation of the lower halides of gallium and indium has already been described in a previous publication (Antell 1959). The object of the present work has been to investigate a method of growing larger crystals of GaP and GaAs using the iodides of gallium. The main interest has been in the growth of GaP, since, with its very high melting point above 1300° C and a dissociation vapour pressure of phosphorus of many atmospheres over the melt, the more orthodox methods of growing crystals by pulling from the melt, etc., raise very serious problems.* Growth of GaP from the vapour phase by the method described here can occur at 900° C and the pressures involved are not much more than 1 atmosphere.

2. Basic reaction

The growth of crystals of GaP from the vapour phase is based on the following overall equation:

$$3GaI + 2P \rightleftharpoons GaI_3 + 2GaP. \qquad (1)$$

As the temperature is raised the reaction tends to proceed from right to left until a new equilibrium is attained. In other words, as the temperature of the reaction rises some of the solid GaP is transferred to the vapour phase. Conversely, if the temperature of the reactants is lowered then crystals of GaP are deposited.

If a reaction tube containing the reactants on the right-hand side of Eqn (1) is placed in a temperature gradient with the GaP at the hotter end, it is found that crystals of GaP grow at the cool end of the tube. This process may be understood as follows. At the hot end of the reaction tube the proportion of GaI in the vapour increases by the absorption of some of the solid GaP. This vapour moves to the cool end of the tube by diffusion and convection and on cooling has an excess of GaI over the equilibrium amount. The vapours can come to equilibrium at the lower temperatures by depositing crystals of GaP as in Eqn (1). The vapour at the cool end then moves to the hot end where it absorbs some more solid GaP. This process will continue until all the solid GaP at the hot end has been transferred to the cool end of the reaction tube. Thus by means of this cyclic process it is possible to transport large amounts of GaP down a temperature gradient by means of a relatively small amount of gallium iodide.

It is often more convenient to enclose the GaP in a reaction tube together with a small amount of iodine in preference to gallium tri-iodide as used in Eqn (1). The initial reaction at the hot end of the reaction tube will then be

$$3GaP + 3I = 3GaI + 3P. \qquad (2)$$

At the cool end of the tube the products on the right-hand side of Eqn (2) react according to Eqn (1). It should be noted that the initial reaction in Eqn (2) gives rise to three atoms of phosphorus whereas only two are required for Eqn (1). Thus it would be expected that the pressure of phosphorus in the reaction tube would be higher, starting with Eqn (2) rather than Eqn (1) provided that the same number of iodine atoms were added in each case. An intermediate pressure of phosphorus could be obtained by adding the iodine atoms partly as GaI_3. This result may be of importance if it proves necessary to grow crystals under a definite pressure of phosphorus in order to obtain stoichiometry in the GaP.

3. Growth of crystals

When a simple reaction tube containing GaP or GaAs and a small quantity of iodine is placed in a temperature gradient a large number of small crystals grow at the cool end of the tube. The size of the crystals could be increased by using a larger reaction tube, a larger quantity of material and by adjusting the temperature gradient by trial and error to give a slow rate of growth. A better solution appeared to be to create a localized cold spot on a relatively small area of the wall of the reaction tube and thereby force the crystals to grow in this region. Such an arrangement was realized by fixing a side arm of about 1·5 cm bore into the furnace tube as shown in Fig. 1. The loss of radiation through the side arm created a cool region in the furnace. By looking down the side arm the growth of the crystals could be easily observed.

The furnace had three windings. The outer two were

* Since this work was written it has been announced (Frosch 1961) that GaP has been grown by a pressurized floating zone technique.

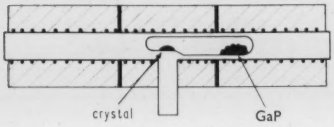


Fig. 1. Section through crystal growing furnace.

connected in series and the centre winding was controlled separately. The temperature distributions along the furnace tube are shown in Fig. 2 curve 1, with the centre furnace switched off and curve 2 with the centre furnace set at a temperature that allowed a crystal of GaP to grow slowly. The dotted portions of each curve represent the unknown temperature of the cold spot.

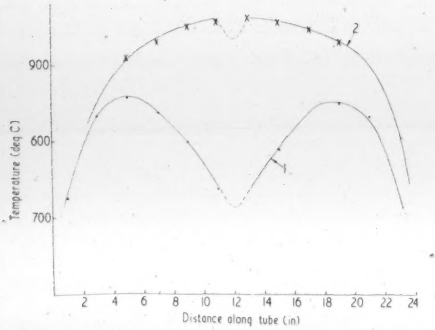


Fig. 2. Temperature distribution in crystal growing furnace.

To grow a crystal, some fragments of GaP were placed in a silica tube about 15 cm long and 2·5 cm diameter together with sufficient purified iodine to give a concentration of about 1·2 mg/cm^3 in the reaction tube. The concentration of iodine does not appear to be critical since concentrations ranging from 0·31 to 3·37 mg/cm^3 have been successfully employed. The reaction tube was placed in the furnace with the GaP at one end as shown in Fig. 1. The part of the reaction tube that was opposite the side arm in the furnace was flattened so that the crystals would have a flat base.

The side furnaces were run at a constant voltage and the centre furnace was set hot enough to prevent any crystals from growing on the cold spot. Crystal nuclei were formed on the cold spot by touching the reaction tube with a thin metal rod. Provided the temperature of the centre furnace was high enough these nuclei evaporated fairly rapidly when the metal rod was removed. When the process was repeated with the centre furnace at a lower temperature the rate of evaporation of the nuclei was greatly reduced. It was then relatively easy to adjust the temperature of the centre furnace so that only one or two of these nuclei remained. The centre furnace was then cooled slowly over a period of several days during which time all the GaP at the hot end of the reaction tube was deposited on the cold spot.

If the cooling rate was too fast or very uneven, voids developed in the crystals as shown in Fig. 3, where the voids have been exposed by grinding away the top surface of the crystal. These voids had smoothn curved surfaces and there was no evidence of the flat crystal faces which might have been expected had they been created by the union of independent overgrowths on a rapidly growing crystal.

Fig. 3. Voids in crystal of GaP × 12.

It was found that GaP crystals grown from the vapour phase adhered very strongly to the walls of the reaction tube and that the crystals always cracked on cooling, usually tearing a small section of quartz from the wall. Such cracks are shown in a polycrystalline sample in Fig. 4.

Fig. 4. Polycrystalline GaP showing cracks.

Several methods were tried in attempts to prevent or alleviate the results of sticking and the most successful solution was to deposit an almost invisible film of carbon on the inside of the reaction tube by the pyrolytic reduction of acetone vapours. If the film was too thick, although still transparent, the radiation from it when hot made it very difficult to observe the formation of nuclei. Figs 5 and 6 show what was substantially a single crystal of GaP grown

on quartz which was covered with a very thin film of carbon. During the attempts at establishing a single crystal nucleus some vapour condensed on the carbon film and may have removed it locally. The GaP crystal adhered so strongly to the quartz in this limited region that it broke into several pieces when being removed from the reaction tube. X-ray analysis revealed that in each fragment of the crystal the (110) face was parallel to the wall of the reaction tube.

Fig. 5. GaP deposit top surface.

Fig. 6. GaP deposit bottom surface.

An experiment was carried out to see if it was possible to deposite GaP on to a seed crystal placed in the reaction tube. The furnace was placed with the side arm pointing vertically downwards. A small seed crystal of GaP was placed in the region of the cold spot in a reaction tube containing iodine and GaP as before. The temperature of the cold spot, initially such that the seed crystal began to evaporate slowly, was then lowered with the intention of reaching a temperature at which GaP would begin to be deposited on the seed. The seed was observed to decrease in size and eventually disappear, while at the same time several small crystals began to grow on the cold spot. This result is not unexpected since the seed crystal was not in good thermal contact with the cold spot on the wall of the reaction tube and would therefore be at a higher temperature than the seeds which grew spontaneously on the cold spot.

If the GaP from which crystals are to be grown contains any free gallium then this gallium is deposited on the cold spot and completely inhibits the transfer of GaP.

One polycrystalline sample of GaAs was grown using about 1·2 mg of iodine per cm^3 in the reaction tube. Growth took place rapidly with the temperature of the cold spot less than 800° C. It was found that GaAs grown in this manner did not stick to clean quartz in the way that GaP did. The crystal was found to be p-type although the original GaAs was n-type. This conversion may be due to the pick-up of copper from the quartz (Edmond 1960).

4. Measurements

4.1. *Iodine content of GaP crystals*

A sample of GaP taken from a crystal which had been grown from a vapour containing 7×10^{18} atoms of iodine per cm^3 was analysed by the Associated Electrical Industries mass spectrometer M.S.7 and iodine was not detected, the estimated limit of detection being less than 2 in 10^8. The covalent radius of iodine is 1·33 Å but since the radius of the largest impurity that would just fit interstitially in the GaP lattice is 1·28 Å the very low solubility of iodine is not unexpected.

4.2. *Thermal conductivity of GaP*

The thermal conductivity of a polycrystalline sample 8 mm × 5 mm × 2 mm has been measured. Difficulty was found in making good contact to the sample but the mean of several determinations gave the thermal conductivity as 0·13 cal s^{-1} cm^{-1} deg^{-1} at 300° K. Since the resistivity of the sample was very high this value is the lattice thermal conductivity of GaP.

4.3. *Electrical measurements*

A single crystal sample of GaAs had a Hall coefficient of +54 cm^3 $coulomb^{-1}$, a conductivity of 5·65 ohm^{-1} cm^{-1} and thus a Hall mobility of 305 cm^2 v^{-1} s^{-1}.

All samples of GaP that were grown from the vapour had a very high bulk resistivity. An attempt was made to determine the resistivity of GaP using a rectangular rod of single crystal material 1 mm long. End contacts were formed by fusing on indium at about 600° C. Two small dots of indium about 2 mm apart were also fused on to one face of the rod to act as potential contacts. The potential between the two indium dots was measured by means of a Vibron vibrating reed amplifier and the current through the sample was determined by similarly measuring the potential difference across a series resistor of 850 MΩ. Leakage currents were reduced by using P.T.F.E. insulation. The resistivity of the sample was found to be in the region of 5×10^9 ohm cm.

5. Conclusion

It has been found that by creating a localized cold spot on a reaction tube it is possible to deposit large crystals of GaP and GaAs. The tendency of GaP to adhere strongly to the quartz of the reaction tube can be overcome by masking the quartz with a very thin film of carbon. It is considered that single crystals of GaP and GaAs could be grown reliably from the vapour by controlling the temperature of the side windings of the furnace and also by decreasing the temperature of the cold spot in a smooth manner rather than the stepped method that was used in the above experiments. The contamination of the GaP crystals by iodine from the vapour has been shown to be negligible but the very high resistivities of all the samples of GaP shows that impurities which act as deep traps have been incorporated in the lattice.

The inhibition of the crystal growing process by free gallium has been established.

The low temperature at which GaAs can be grown from the vapour may be useful as a means of incorporating large concentrations of impurities which exhibit a retrograde solubility; that is, impurities whose solubilities are less at the melting point of GaAs than at lower temperatures.

Acknowledgments

The author is indebted to Mr. D. Effer for preparing the GaP powder, to Mr. L. D. Brownlee and Miss K. Whitehead for carrying out the x-ray analysis, and to Miss A. D. Stuckes and Mrs. B. Brooks for the thermal conductivity measurements.

The author would like to thank Mr. R. P. Chasmar for many useful discussions and also for his comments on the manuscript, and to thank Sir Willis Jackson, Director of research, A.E.I. (Manchester) Ltd., for permission to publish.

References

ANTELL, G. R., and EFFER, D., 1959, *J. Electrochem. Soc.*, **106**, 509.

EDMOND, J. T., 1960, *J. Appl. Phys.*, **31**, 1428.

FROSCH, C. J., and DERIK, L., 1961, *J. Electrochem. Soc.*, **108**, 251.

Notes and comments

Journal of Catalysis

We have received notice of the new *Journal of Catalysis*, which is to be published by Academic Press, 111 Fifth Avenue, New York 3. It is to be a medium of publication for all original scientific studies in heterogeneous catalysis, as well as subjects related to homogeneous catalysis. Studies relating catalytic properties with chemical processes at surfaces, including the study of surfaces by various physical means, will be included in the scope of this journal, as will molecular rate processes and studies of the chemistry of surfaces.

The Editors will be Professor J. H. de Boer, Technological University of Delft and Professor P. W. Selwood, Northwestern University. The first issue is expected to appear early in 1962.

International Conference on Magnetic and Electric Resonance and Relaxation

An International Conference on Magnetic and Electric Resonance and Relaxation will be held under the auspices of the Netherlands Physical Society and the Groupement AMPERE from 2nd–7th July, 1962, at Eindhoven, The Netherlands. The subject will be mainly double resonance experiments but also nuclear magnetic resonance, electron spin resonance, dielectric relaxation, quadrupole resonance and related subjects. Abstracts must be submitted by 1st January, 1962.

Further information may be obtained from the secretary of the committee, Dr. D. J. Kroon, Philips Research Laboratories, Eindhoven, The Netherlands.

Diffusion in heterogeneous media: lattices of parallelepipeds in a continuous phase

by R. M. BARRER, Sc.D., F.R.S., and J. H. PETROPOULOS, Ph.D., Physical Chemistry Laboratories, Chemistry Department, Imperial College, London, S.W.7

MS. received 27th July 1961

Abstract

A treatment has been given of the permeability of a heterogeneous medium consisting of a regular lattice of rectangular parallelepipeds in a continuum. Effects on the permeability due to surface resistances on the upstream and downstream faces of the parallelepipeds have been considered, and also aspects of the role of particle shape and orientation. The treatment provides an approach to the permeability of and fractionation by foam plastic membranes. When the disperse phase has molecular sieve properties, high continuous fractionation factors for appropriate composite membranes and mixtures of diffusants are found to be possible. The treatment is also valid for the thermal or electrical conductivity, permittivity, or magnetic permeability of such a lattice.

1. Introduction

CALCULATIONS of diffusion fluxes through heterogeneous media, or of permittivity or electrical or thermal conductivity in such media, are all aspects of a single problem. A usual heterogeneous membrane consists of a phase A dispersed as small randomly arranged and often randomly shaped particles in a continuum of phase B. Of the various approximations which have been made to dielectric, electrical, thermal or diffusion behaviour in such a system the following are the most important.

Maxwell (1873) considered a continuum with immersed spheres so far apart that the streamline pattern about each sphere was uninfluenced by those of its neighbours. Maxwell's equation is thus valid for low volume fractions of disperse phase. Equations appropriate to higher volume fractions of disperse phase have been developed for dielectrics by assuming each sphere to be surrounded by an essentially uniform medium of dielectric constant equal to the overall dielectric constant for the composite medium. The most useful and successful relations of this type are those of Bruggeman (1935) and of Böttcher (1945), although the assumption made in their derivation led Poley (1953) to question their general validity. In the same direction, Niesel (1952, 1953) developed, and Eichbaum (1959) studied experimentally, formulae for the dielectric constant of media comprising spheres, and randomly oriented lamellae and needles embedded in continua; while Sillars (1937) and van Beek (1960) considered embedded spheroids, oblate and prolate. Brown (1955) treated the problem statistically and arrived at a series solution, the first two terms of which were independent of particle shape and arrangement and were equivalent to Maxwell's relation. By assigning particular values to a structure factor in the third term, the solution can be made to agree with either the Bruggeman or the Böttcher formula. Prager (1960) has also applied statistical methods to the problem.

Another development of Maxwell's formula is due to Rayleigh (1892) who, using potential theory, calculated the influence on the flow pattern around any one sphere due to neighbouring spheres forming a regular cubic array. This treatment was re-developed by Runge (1925) and extended by Meredith and Tobias (1960) who also provided some experimental verification for spheres of electrical conductivity either very high or very low in comparison with that of the surrounding medium. Similar verification has been given by Klute (1959) for the latter case. Little is known, however, about the applicability of the formula if the disperse phase is present as particles of random shape, size and arrangement. Formulae have also been given (Rayleigh 1892, Runge 1925) for regular lattices of cylindrical rods or thin tubes, extending throughout the breadth of a heterogeneous membrane, parallel to each other and normal to the direction of flow.

In the present paper, in connection with an experimental programme on properties of heterogeneous membranes, we consider the permeability of a membrane in which a regular array of rectangular parallelepipeds of phase A is embedded in a continuum of B. A treatment in terms of potential theory would be difficult here, and an alternative phenomenological procedure has been developed. Since the model and procedure permit consideration of aspects largely complementary to those that can be examined by the other methods, it was thought of interest to give an account of this procedure and the results to which it leads.

2. The permeability of the model

The disperse phase A is in the form of rectangular parallelepipeds of square cross section of side a, and length h_A along the direction of flow, disposed regularly in a continuous matrix of phase B at distances apart h_B, in the direction of the flow, and b at right angles to that direction, as shown in Fig. 1.

The diffusion coefficients D_A and D_B of a diffusant in A and B are independent of concentration and the distribution of diffusant between phases A and B at equilibrium is governed by Henry's law, i.e.

$$\left(\frac{C_A}{C_B}\right)_{\text{equil.}} = k = \text{const.}$$

Here C_A, C_B denote the concentration of diffusant in phases A, B respectively (as number of molecules per unit volume of that phase).

The type of streamline pattern established in steady state diffusion depends on the relation between kD_A and D_B (i.e. the relative permeabilities of the component phases). If $kD_A = D_B$ the composite medium behaves as a homogeneous one and all streamlines are parallel.

In the general case $kD_A \neq D_B$ and there will be convergence and divergence of streamlines, a higher density of streamlines being established in the more permeable component phase.

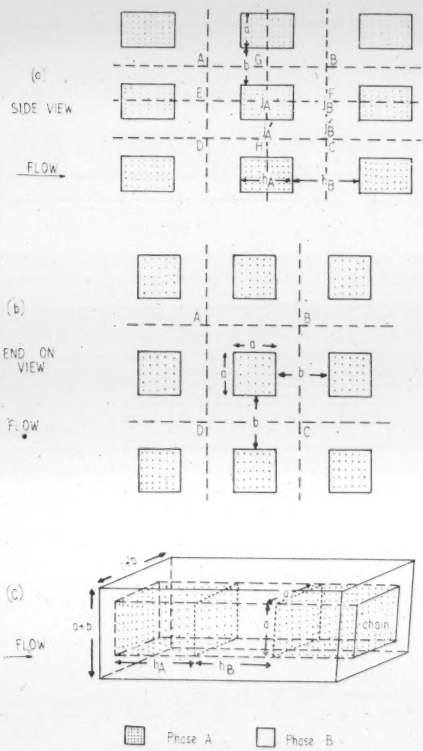


Fig. 1.

Because of symmetry, however, flow lines along planes parallel to the direction of flow and either bisecting cells of phase A or equidistant between such cells (AB, CD and EF in Fig. 1(*a*)) remain undisturbed. Similarly, all streamlines will be parallel to the direction of flow in planes normal to this direction and either bisecting, or equidistant between, cells of phase A (AD, BC and GH in Fig. 1(*a*)). Hence the flow pattern is repeated in every region of the lattice, such as that enclosed by ABCD in Figs. 1(*a*) and (*b*) which, therefore, represents the 'unit cell' of the composite medium.

Consider an array of such unit cells of length l equal to the thickness of the heterogeneous membrane (a portion of such an array is shown in Fig. 1(*c*)). This array constitutes a chain of cross-sectional area a^2 consisting of alternate cells of phases A and B of length h_A and h_B respectively, surrounded by phase B in the form of a hollow tube of square cross section of area $b(2a + b)$.

For the steady state of flow, and confining attention first to the chain, fluxes j_A, j_B may be defined respectively through the mid-sections of cells of A and B. These fluxes are not equal because of the non-parallelism of flow lines. Hence we may write

$$j_A = \gamma j_B. \tag{1}$$

The concentration of diffusant at interfaces normal to the direction of flow will be denoted by C with subscripts A and B to indicate the side of the interface for which the concentration is defined, and subscripts 1 and 2 to denote in-going and out-going faces respectively. The concentrations so defined are mean values since the concentration of diffusant varies from point to point on an interface.

It is also convenient to define 'mean fluxes' $\bar{j}_A$, $\bar{j}_B$ through each cell of phases A and B respectively, in terms of the concentration difference between in-going and out-going faces, and to relate them to j_A and j_B by means of parameters β_A, β_B respectively:

$$j_A = \beta_A \bar{j}_A = \frac{\beta_A a^2 D_A}{h_A}(C_{A_1} - C_{A_2}) \tag{2}$$

$$j_B = \beta_B \bar{j}_B = \frac{\beta_B a^2 D_B}{h_B}(C_{B_1} - C_{B_2}). \tag{3}$$

Since, in general, there is a possibility that a definite physical barrier is associated with the crossing of an interface, an expression for the flux across an interface j_s may be written in terms of a transmission coefficient ν and the concentrations on either side of the interface and may be related to j_A by means of a parameter α^*. Thus

$$j_s = a^2\nu(kC_{B_2} - C_{A_1}) = \alpha j_A. \tag{4}$$

If l is the length of the chain (i.e. the thickness of the heterogeneous membrane), l, l_A, l_B being the aggregate lengths of cells of phases A and B respectively, and if the chain contains n cells of each phase, then

$$\begin{aligned} l_A &= nh_A,\ l_B = nh_B \\ l &= l_A + l_B = n(h_A + h_B) = nh. \end{aligned} \tag{5}$$

The concentration drop across the composite membrane as measured in phase A is then

$$\begin{aligned} \Delta C_A &= n(C_{A_1} - C_{A_2}) + nk(C_{B_1} - C_{B_2}) \\ &\quad + (2n - 1)(kC_{B_2} - C_{A_1}) \\ &= n(C_{A_1} - C_{A_2}) + nk(C_{B_1} - C_{B_2}) \\ &\quad + 2n(kC_{B_2} - C_{A_1}) \end{aligned} \tag{6}$$

as $n \gg 1$. From Eqns (1) to (4) and (6) we get

$$\Delta C_A = n(C_{A_1} - C_{A_2})\left\{1 + \frac{\beta_A D_A}{h_A}\left(\frac{2\alpha}{\nu} + \frac{h_B k}{\gamma \beta_B D_B}\right)\right\}. \tag{7}$$

The tube of medium B surrounding the chain is next visualized as divided into cells of length h_A and h_B alternately, to correspond with the cells of phases A and B in the chain. All fluxes, concentrations and auxiliary parameters are defined in the same way as above and are denoted by the same symbols primed (except that in this case, of course, it is not necessary to include fluxes across interfaces). Thus, at the midsections of cells surrounding cells of phase A in the chain, a flux j'_A may be defined, and at the midsections of cells surrounding cells of phase B in the chain a flux j'_B.

* If the barriers at upstream and downstream faces are different, we may define parameters ν_1, α_1, ν_2, α_2 for these faces respectively. In the final result (Eqn 16), the term $2\alpha/\nu$ is then replaced by $\alpha_1/\nu_1 + \alpha_2/\nu_2$.

These fluxes are related to each other and to the mean fluxes as follows:

$$j_A = \beta'_A \bar{j}'_A = \frac{\beta'_A b(2a+b)D_B}{h_A}(C'_{A_1} - C'_{A_2}) = \gamma' j'_B$$

$$= \gamma' \beta'_B \bar{j}'_B = \frac{\gamma' \beta'_B b(2a+b)D_B}{h_B}(C'_{B_1} - C'_{B_2}). \quad (8)$$

The concentration drop in the surrounding medium across the composite membrane as measured in phase B is then

$$\Delta C_B = n(C'_{A_1} - C'_{A_2}) + n(C'_{B_1} - C'_{B_2})$$

$$= n(C'_{A_1} - C'_{A_2})\left(1 + \frac{\beta'_A h_B}{\gamma' \beta'_B h_A}\right). \quad (9)$$

Because of the steady state conditions we also have

$$\Delta C_A = k \Delta C_B \quad (10)$$

$$j_A + j'_A = j_B + j'_B; \quad \frac{j_B}{j'_B} = \frac{1-\gamma'}{\gamma - 1}. \quad (11)$$

The overall flux J through the array of unit cells of the composite membrane (Fig. 1(*c*)) is equal to the sum of the fluxes through the chain and the surrounding medium. Thus the overall diffusion coefficient $\bar{D}$, characteristic of the whole composite medium is given by

$$J = (a+b)^2 \bar{D} \frac{\overline{\Delta C}}{l} \quad (12)$$

where $\overline{\Delta C}$ is the combined concentration drop in phases A and B over the thickness of the heterogeneous membrane and is defined by*

$$\overline{\Delta C} = v_A \Delta C_A + v_B \Delta C_B. \quad (13)$$

Here v_A,v_B are the volume fractions occupied by phases A and B respectively in the composite medium. Hence:

$$J = (a+b)^2 \overline{D}(v_A \Delta C_A + v_B \Delta C_B)/l. \quad (14)$$

Also, as pointed out above, we have

$$J = j_A + j'_A. \quad (15)$$

Hence, from Eqns (2), (8), (14), (15), and using Eqn (10), we finally obtain:

$$\bar{D} = \frac{h}{(a+b)^2(kv_A + v_B)} \left\{ \frac{ka^2 \beta_A D_A}{h_A + \beta_A D_A\left(\frac{2\alpha}{\nu} + \frac{h_B k}{\gamma \beta_A D_B}\right)} + \frac{b(2a+b)\beta'_A D_B}{h_A + \frac{\beta'_A h_B}{\gamma' \beta'_B}} \right\} \quad (16)$$

where the h's may be substituted by l's from Eqn (5).

The distance parameters of Eqn (16) are related to v_A, v_B as follows:

$$v_A = \frac{a^2 l_A}{(a+b)^2 l} \text{ (with } v_B = 1 - v_A,\ l_B = l - l_A). \quad (17)$$

Thus Eqn (16) becomes

$$\bar{D} = \frac{l}{l_A(kv_A + v_B)} \left\{ \frac{lkv_A \beta_A D_A}{l_A + \beta_A D_A\left(\frac{2n\alpha}{\nu} + \frac{l_B k}{\gamma \beta_B D_B}\right)} + \frac{(l_A - v_A l)\beta'_A D_B}{l_A + \frac{\beta'_A l_B}{\gamma' \beta'_B}} \right\}. \quad (18)$$

* One may also consider only the concentration drop ΔC_B in the continuous phase, i.e. $\Delta C = \Delta C_B$ in place of Eqn (13). The term $(kv_A + v_B)$ is then omitted from all formulae for D to give the relevant modified diffusion coefficients. For example, Eqn 36 becomes $D_2 = (1 - v_A^{2/3})D_B$.

For a 'cubic lattice' there are the additional relations:

$$a = h_A;\ b = h_B;\ \frac{l_A}{l} = v_A^{\frac{1}{3}};\ \frac{l_B}{l} = 1 - v_A^{\frac{1}{3}} \quad (19)$$

and the expression for $\bar{D}$ becomes

$$\bar{D} = \frac{1}{kv_A + v_B} \left[\frac{kv_A^{\frac{1}{3}} \beta_A D_A}{v_A^{\frac{1}{3}} + \beta_A D_A\left\{\frac{2n\alpha}{\nu} + \frac{(1 - v_A^{\frac{1}{3}})k}{\gamma \beta_B D_B}\right\}} + \frac{(1 - v_A^{\frac{1}{3}})\beta'_A D_B}{v_A^{\frac{1}{3}} + \frac{(1 - v_A^{\frac{1}{3}})\beta'_A}{\gamma' \beta'_B}} \right]. \quad (20)$$

In the absence of surface diffusion barriers, the term $2\alpha/\nu$ in Eqn (16) and the term $2n\alpha/\nu$ in Eqn (18) and Eqn (20) vanish.

The values of the β and γ parameters of Eqns (16), (18) and (20) would, in general, be expected to depend on composition. The following inequalities apply to these parameters:

(i) if $kD_A > D_B$, then $\gamma > 1$, $\gamma' < 1$, $\beta_A > 1$, $\beta_B < 1$, $\beta'_A < 1$, $\beta'_B > 1$;

(ii) if $kD_A < D_B$, then $\gamma < 1$, $\gamma' > 1$, $\beta_A < 1$, $\beta_B > 1$, $\beta'_A > 1$, $\beta'_B < 1$.

Thus in the complex coefficients $\beta_A/\gamma\beta_B$ and $\beta'_A/\gamma'\beta'_B$ there is always partial compensation between individual factors.

3. Special cases

3.1. *Impermeable fillers*

The equation appropriate to impermeable fillers is obtained when $kD_A = 0$. In this case the only equations relevant are Eqns (8) and (9). In Eqn (11) $j_A = 0$, $\gamma = 0$, and the equation becomes

$$j'_A = j_B + j'_B;\ j_B/j'_B = \gamma' - 1. \quad (21)$$

Two cases may be distinguished:

(i) Phase A is non-conducting ($D_A \to 0$; k is finite). In this case Eqn (14) holds, and Eqn (15) becomes

$$J = j'_A. \quad (22)$$

Finally, Eqn (16) becomes

$$\bar{D} = \frac{hb(2a+b)\beta'_A D_B}{(a+b)^2(kv_A + v_B)h_A[1 + (\beta'_A h_B/\gamma' \beta'_B h_A)]}. \quad (23)$$

(ii) Phase A is non-sorbing ($k \to 0$, D_A is finite). Eqn (22) also holds in this case but $\Delta C_A = 0$ in Eqn (14), so that the final expression for $\bar{D}$ is given by Eqn (23) with $kv_A = 0$. The formula appropriate to a lattice of cubes, analogous to Eqn (20), is

$$\bar{D} = \frac{1 - v_A}{1 - v_A}\left\{\frac{\beta'_A D_B}{v_A^{\frac{1}{3}} + \beta'_A(1 - v_A^{\frac{1}{3}})/\gamma' \beta'_B}\right\}. \quad (24)$$

Under the conditions considered here there is little deviation of stream-lines from parallelism in the part of the medium in which fluxes j'_A, $\bar{j}'_A$ occur, i.e., $\beta'_A \simeq 1$ in Eqns (23) and (24). Eqn (24) can then be rewritten in the form

$$\frac{D_B(1 - v_A^{\frac{1}{3}})}{\bar{D}(1 - v_A)(1 - v_A^{\frac{1}{3}})} - \frac{v_A^{\frac{1}{3}}}{1 - v_A^{\frac{1}{3}}} = \frac{1}{\gamma' \beta'_B} \quad (24a)$$

giving the variation of the coefficient $1/\gamma'\beta'_B$ with v_A in terms of D_B and $\bar{D}$.

3.2. *Particular arrangements of filler*

3.2.1. *Phase A as thin laminae normal to the direction of flow* (h_A *small*, $h_B \to h$). The condition $h_A \ll h_B$ ($h_B \to h$) also implies that $v_A \ll v_B$, $v_B \to 1$ and, if k is not unduly large so that $kv_A \ll v_B$, Eqn (16) reduces to

$$\bar{D} = \frac{D_B}{(a+b)^2}\left\{\frac{a^2}{\frac{2n\alpha}{vlk}D_B + \frac{1}{\gamma\beta_B}} + b(2a+b)\gamma'\beta_B'\right\}. \quad (25)$$

Eqn (25) gives the diffusion coefficient when the otherwise continuous medium B contains a series of discontinuities acting as diffusion barriers, such as grain boundaries or films in grain boundaries, perpendicular to the direction of flow.

When the boundaries extend in almost unbroken continuity across the medium, $b \ll a$ and Eqn (25) reduces to

$$\bar{D} = D_B \Big/ \left(\frac{2n\alpha}{vlk}D_B + 1\right) \quad (26)$$

or if $2n\alpha/vlk \gg 1/D_B$ (i.e. the diffusion is dominated by the surface barriers)

$$\bar{D} = \frac{vlk}{2n\alpha}. \quad (27)$$

In the absence of surface barriers, on the other hand, Eqn (25) becomes

$$\bar{D} = \frac{D_B}{(a+b)^2}\left\{a^2\gamma\beta_B + b(2a+b)\gamma'\beta_B'\right\}. \quad (28)$$

If $kD_A \gg D_B$ then γ, γ', β_B', $\beta_B \to 1$ and the above equation yields $\bar{D} = D_B$, as expected. The equation for impermeable thin lamellae ($D_A = 0$, $\gamma = 0$), on the other hand, is

$$\bar{D} = \frac{D_B b(2a+b)\gamma'\beta_B'}{(a+b)^2} \quad (29)$$

which reduces, as required, to $\bar{D} = 0$ for continuous boundaries ($b = 0$).

The results of this section are in certain respects an extension of the treatment of Ash and Barrer (1959) who discussed permeability and diffusion in a homogeneous membrane with a surface resistance at the faces of entry and exit only. Some possibilities are now seen for internal surface resistances distributed within the body of the medium and not necessarily continuous over its whole cross-sectional area. The measured diffusion coefficients become structure sensitive and less than the true diffusion coefficient D_B, while fluxes may no longer be inversely proportional to thickness. Some of the loss of permeability often observed in H_2–Pd diffusion membranes after use (e.g. Barrer 1940) may arise from the opening of physical discontinuities within the structure. Such thin fissures, formed by the pressure of hydrogen molecules discharged as atom pairs from the bulk metal into grain boundaries, can act as resistances to transport of atomic hydrogen. Many such resistances must be transverse to the direction of flow. Mechanical weakness after absorption of hydrogen is known for iron as well as palladium, and could also be due to the same cause.

3.2.2. *Laminated media.* In such media both phases A and B are continuous in two dimensions, but alternate in the third.

Two cases may be distinguished according to whether flow occurs along or at right angles to the laminae. In both cases a parallel flow-line pattern is obtained so that all β's and γ's are equal to unity.

(i) Flow at right angles to the laminae. In this case $b \to 0$ in Eqn (16), and also $l_A/l = v_A$, $l_B/l = v_B$. The equation then reduces to

$$\bar{D} = \frac{1}{(kv_A + v_B)}\left[\frac{kD_A}{v_A + D_A\left(\frac{2n\alpha}{lv} + \frac{v_B k}{D_B}\right)}\right] \quad (30)$$

which when surface barriers dominate, becomes

$$\bar{D} = \frac{lk}{kv_A + v_B}\,\frac{v}{2n\alpha}. \quad (31)$$

(ii) Flow parallel to the laminae. In this case $h_B \to 0$, $h_A \to h$. Also

$$v_A = a^2/(a+b)^2; \quad v_B = b(2a+b)/(a+b)^2.$$

Further, there are no phase interfaces at right angles to the direction of flow so that the surface barrier term in Eqn (16) may be omitted. Eqn (16) then reduces to

$$\bar{D} = \frac{kv_A D_A + v_B D_B}{kv_A + v_B}. \quad (32)$$

The effect of orientation may be considered by comparing Eqns (30) and (32). Thus, in the absence of surface barriers, the ratio of diffusion coefficients for flow parallel and perpendicular to the laminae respectively is

$$\frac{\bar{D}_{\parallel}}{\bar{D}_{\perp}} = \left(v_A + \frac{v_B D_B}{kD_A}\right)\left(v_A + \frac{kD_A v_B}{D_B}\right). \quad (33)$$

Values of this ratio for various values of kD_A/D_B and v_A are given in Table 1, from which major influences of orientation of lamellae upon diffusion are seen.

Table 1. *Diffusion anisotropy associated with laminae differently oriented to the flow direction.*

$D_B = 10^{-8}$ cm^2/s $\quad k = 1$

D_A (cm^2/s)	kD_A/D_B	$\bar{D}_{\parallel}(v_A = \frac{1}{2})$ (cm^2/s)	$\bar{D}_{\perp}(v_A = \frac{1}{2})$ (cm^2/s)	$\bar{D}_{\parallel}/\bar{D}_{\perp}$ $v_A = \frac{1}{2}$	$v_A = \frac{1}{3}$	$v_A = \frac{1}{4}$	$v_A = \frac{1}{5}$
10^{-5}	1000	$5{\cdot}005 \times 10^{-5}$	$1{\cdot}99 \times 10^{-7}$	251	223	189	161
10^{-6}	100	$5{\cdot}05 \times 10^{-6}$	$1{\cdot}98 \times 10^{-7}$	25·5	22·8	19·4	16·7
10^{-7}	10	$5{\cdot}5 \times 10^{-7}$	$1{\cdot}82 \times 10^{-7}$	3·03	2·80	2·53	2·30
5×10^{-8}	5	$3{\cdot}0 \times 10^{-8}$	$1{\cdot}67 \times 10^{-8}$	1·80	1·72	1·60	1·51
3×10^{-8}	3	$2{\cdot}0 \times 10^{-8}$	$1{\cdot}5 \times 10^{-8}$	1·33	1·30	1·25	1·21
2×10^{-8}	2	$1{\cdot}5 \times 10^{-8}$	$1{\cdot}33 \times 10^{-8}$	1·13	1·11	1·09	1·08
1×10^{-8}	1	$1{\cdot}0 \times 10^{-8}$	$1{\cdot}0 \times 10^{-8}$	1·00	1·00	1·00	1·00
5×10^{-9}	$\frac{1}{2}$	$7{\cdot}5 \times 10^{-9}$	$6{\cdot}63 \times 10^{-9}$	1·13	1·11	1·09	1·08
$3{\cdot}3 \times 10^{-9}$	$\frac{1}{3}$	$6{\cdot}65 \times 10^{-9}$	$5{\cdot}0 \times 10^{-9}$	1·33	1·30	1·24	1·21
$2{\cdot}0 \times 10^{-9}$	$\frac{1}{5}$	$6{\cdot}0 \times 10^{-9}$	$3{\cdot}33 \times 10^{-9}$	1·80	1·72	1·60	1·51
10^{-9}	$\frac{1}{10}$	$5{\cdot}5 \times 10^{-9}$	$1{\cdot}81 \times 10^{-9}$	3·03	2·80	2·53	2·30
10^{-10}	$\frac{1}{100}$	$5{\cdot}05 \times 10^{-9}$	$1{\cdot}98 \times 10^{-10}$	25·5	22·8	19·4	16·7
10^{-11}	$\frac{1}{1000}$	$5{\cdot}0 \times 10^{-9}$	$1{\cdot}99 \times 10^{-11}$	251	223	189	161

3.3. *Media in which the disperse phase occupies a large volume fraction*

(i) $kD_A > D_B$. Provided that v_A, kD_A are both considerably larger than v_B, D_B respectively, then the flux carried by the chain will be much greater than that carried by the surrounding medium and the disturbance of the flow line pattern from parallelism in the chain will be a relatively minor effect. In particular we may write $\beta_A \simeq 1$, in view of the fact that $j_A \gg j'_A$. Fluxes j_B, j'_B per unit cross-sectional area are comparable, however, and in this case it is more accurate to write $j_B/a^2 = j'_B/b(2a+b)$. This implies in effect that $C_{B_1} = C'_{B_1}$, $C_{B_2} = C'_{B_2}$. The best way of approximating to j'_A seems to be to assume $\beta'_A = 1$ and $C_{A_1} = C'_{A_1}$, $C_{A_2} = C_{A_2}$ (since in any case we have $j_A \gg j'_A$). By the methods of the present paper it is possible to arrive at the following expression for the 'cubic model':

$$\bar{D}_1 = \frac{1}{kv_A + v_B}\left\{\frac{1 - v_A^{\frac{1}{3}}}{D_B} + \frac{v_A^{\frac{1}{3}}}{v_A^{\frac{1}{3}}kD_A + (1 - v_A^{\frac{1}{3}})D_B}\right\}^{-1}. \quad (34)$$

For sufficiently high values of v_A and kD_A/D_B, j'_A is negligible and Eqn (34) reduces to

$$\bar{D}_1 = \frac{1}{kv_A + v_B}\left\{\frac{1 - v_A^{\frac{1}{3}}}{D_B} + \frac{1}{v_A^{\frac{1}{3}}kD_A}\right\}^{-1}. \quad (35)$$

Eqn (34) is also valid when $kD_A \simeq D_B$, irrespective of the value of v_A.

Eqns (34) and (35) are of interest for the description of diffusion in foams and in foam plastics containing a large volume fraction of gas gaps and also in polymers containing highly permeable filler particles, provided these can be packed sufficiently closely to ensure a high v_A value.

(ii) $kD_A \to 0$. If phase A is very nearly impermeable, then, provided v_A is sufficiently large, a simple expression for the diffusion coefficient of the cubic model of Fig. 1 may be obtained by disregarding the small modification of the flux caused by the blind pores, i.e. by putting $\beta_A = \beta'_B = \gamma' = 1$ in Eqn (24)

$$\bar{D}_2 = \frac{1 - v_A^{\frac{1}{3}}}{v_B}D_B. \quad (36)$$

4. Separation of diffusants

Consider two diffusants in a composite membrane so chosen that phase A is a molecular sieve permeable to the first (subscript 1) but not to the second (subscript 2). Then for large v_A and kD_{A_1}/D_{B_1}, assuming that ΔC_B is the same for both species, the ratio of the fluxes may be obtained from Eqns (34) and (36):

$$\frac{J_1}{J_2} = \frac{1}{1 - v_A^{\frac{1}{3}}}\left\{\frac{(1 - v_A^{\frac{1}{3}})D_{B_2}}{D_{B_1}} + \frac{v_A^{\frac{1}{3}}D_{B_2}}{v_A^{\frac{1}{3}}kD_{A_1} + (1 - v_A^{\frac{1}{3}})D_{B_1}}\right\}^{-1}. \quad (37)$$

A simpler expression is obtained if Eqn (35) is employed instead of Eqn (34), namely:

$$\frac{J_1}{J_2} = \frac{1}{1 - v_A^{\frac{1}{3}}}\left\{\frac{(1 - v_A^{\frac{1}{3}})D_{B_2}}{D_{B_1}} + \frac{1}{v_A^{\frac{1}{3}}}\frac{D_{B_2}}{kD_{A_1}}\right\}^{-1}. \quad (38)$$

For sufficiently high values of kD_{A_1}/D_{B_2} and v_A there is very little difference between Eqns (37) and (38) as may be seen from Table 2. In this table the value of the flux ratio obtained from these equations (the figures obtained from Eqn (37) being in parentheses) are listed for the case $D_{B_1} = D_{B_2} = D_B$ ($D_{A_1} = D_A$) and for various values of kD_A/D_B and v_A. Evidently very large and continuous fractionation factors are possible with molecular sieve crystallites embedded in another medium and occupying substantial volume fractions of the whole.

Table 2. *Flux ratios J_1/J_2 based on Eqns* (37) *and* (38). *The figures in parentheses relate to Eqn* (37)

v_A	$\frac{kD_A}{D_B} = 1$		$\frac{kD_A}{D_B} = 10$		$\frac{kD_A}{D_B} = 100$		$\frac{kD_A}{D_B} = 1000$		$\frac{kD_A}{D_B} = \infty$	
0·60	2·58	(3·46)	12·6	(12·8)	20·6	(20·6)	22·0	(22·0)	22·1	(22·1)
0·65	3·11	(4·02)	16·1	(16·3)	27·6	(27·6)	29·7	(29·7)	29·9	(30·0)
0·70	3·82	(4·72)	21·0	(21·3)	38·3	(38·2)	41·7	(41·7)	42·2	(42·1)
0·75	4·81	(5·73)	28·4	(28·8)	55·9	(55·9)	61·9	(61·9)	62·7	(62·7)
0·80	6·30	(7·23)	40·3	(40·7)	87·7	(87·8)	99·4	(99·4)	101	(101)
0·85	8·79	(9·75)	61·5	(62·0)	154	(154)	181	(181)	185	(185)
0·90	13·8	(14·75)	107	(107)	328	(329)	415	(415)	427	(427)
0·95	28·8	(29·8)	251	(251)	1100	(1100)	1660	(1660)	1760	(1760)

Fluxes through, and separations by, foam plastics are of much interest. Here phase A is permeable to both diffusants. From Eqn (35) it follows that the ratio R of the flux through pure polymer J_p to that through the foam polymer J_f is approximately

$$\frac{J_p}{J_f} = R = \frac{D_B}{kD_A}\frac{1}{v_A^{\frac{1}{3}}} + (1 - v_A^{\frac{1}{3}}). \quad (39)$$

Since in the foam plastic we can easily have $kD_A \gg D_B$, the permeability in this case can be much higher (Table 3(*a*)). Moreover, if R_p denotes the ratio of the fluxes of species 2 and 1 respectively, and R_f is this ratio for the foam plastic then

$$\frac{R_p}{R_f} = \frac{(D_{B_2}/k_2D_{A_2}) + v_A^{\frac{1}{3}} - v_A^{\frac{2}{3}}}{(D_{B_1}/k_1D_{A_1}) + v_A^{\frac{1}{3}} - v_A^{\frac{2}{3}}}. \quad (40)$$

Clearly, for appropriate values of the ratios kD_A/D_B for species 1 and 2 and of the volume fraction of cavities in the foam plastic, a better fractionation factor may be obtained from a foam plastic (Table 3(*b*)). Even when use of such a plastic has a slightly adverse effect on the fractionation, advantage may be gained from the increased fluxes noted above. If kD_A/D_B is sufficiently large for both diffusing species (e.g. $\geqslant 1000$), then Eqn (40) is dominated by the terms in v_A and $R_p/R_f \simeq 1$ for all values of v_A.

The validity of this equation and the advantages in fractionation that it indicates depend on the absence of unduly large bubbles or networks of interconnected channels in the polymer matrix. The difficulty of fulfilling this requirement in practice would be expected to increase with increasing v_A.

In the simultaneous diffusion of two species in laminated media different flux ratios arise according to the direction of flow with respect to the laminations. The ratio for flow normal to the laminations, obtained from Eqn (30) and in the absence of surface barriers to diffusion, is

$$R_\perp = \frac{\Delta C_{B_1}k_1D_{A_1}}{\Delta C_{B_2}k_2D_{A_2}}\frac{1 + (v_Bk_2D_{A_2}/v_AD_{B_2})}{1 + (v_Bk_1D_{A_1}/v_AD_{B_1})}. \quad (41)$$

Table 3(*a*). *Ratio of fluxes J in a foam plastic and the massive polymer*

$\frac{kD_A}{D_B}$	J_f/J_p			
	$v_A = 0{\cdot}85$	$v_A = 0{\cdot}90$	$v_A = 0{\cdot}95$	$v_A = 0{\cdot}98$
10	6·3	7·3	8·4	9·4
20	9·4	11·6	14·7	17·5
100	15·6	22·3	36·8	59·5
1000	18·5	28·2	55·6	129
10000	18·9	28·9	58·8	147

Table 3(*b*). *Relative values of flux ratios R_p and R_f through massive polymer and foam plastic*

$\frac{k_2D_{A_2}}{D_{B_2}}$	$R_p/R_f\left(R_p = \frac{J_{2p}}{J_{1p}}, R_f = \frac{J_2}{J_{1f}}\right)$							
	$v_A = 0{\cdot}85$		$v_A = 0{\cdot}90$		$v_A = 0{\cdot}95$		$v_A = 0{\cdot}98$	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
10	1·19	1·22	3·09	3·91	4·38	6·61	6·39	13·86
20	1·08	1·11	1·93	2·42	2·50	3·77	3·49	7·35
100	1·00	1·02	1·00	1·26	1·00	1·51	1·00	2·17
1000	0·981	1·00	0·79	1·00	0·663	1·00	0·461	1·00
10000	0·978	0·998	0·77	0·97	0·629	0·948	0·407	1·013

Column (1) $k_1D_{A_1}/D_{B_1} = 100$, column (2) $k_1D_{A_1}/D_{B_1} = 1000$.

If the flow is parallel to the laminations, Eqn (32) yields

$$R_{\parallel} = \frac{\Delta C_{B_1}}{\Delta C_{B_2}} \frac{k_1 v_A D_{A_1} + v_B D_{B_1}}{k_2 v_A D_{A2} + v_B D_{B_2}}. \qquad (42)$$

The ratio $R_0 = R_{\perp}/R_{\parallel}$ then measures the anisotropy of the fractionation by flow. Some numerical examples, demonstrating the considerable anisotropy under various conditions, are given in Table 4.

Gabor 1959). However, for the lattices considered in this paper, the value of $\bar{D}$ would be nearly the same in both states of flow provided $kD_A \gg D_B$. Thus, in diffusion through a foam plastic the time lag L will be (Barrer 1951) $L = l^2/6\bar{D}$, where $\bar{D}$ is given approximately by Eqn (34) or (35). For some problems of thermal insulation time lags are as important as steady state fluxes and have been considered by Jaeger (1950) for composite media in the form of several

Table 4. *Anisotropy of flux ratios in a laminated medium with laminae normal and parallel respectively to flow direction*

$\frac{k_1D_{A_1}}{D_{B_1}}$	$\frac{k_1D_{A_2}}{D_{B_2}}$	$R_{\perp}/R_{\parallel}$			$\frac{k_1D_{A_1}}{D_{B_1}}$	$\frac{k_2D_{A_2}}{D_{B_2}}$	$R_{\perp}/R_{\parallel}$
		$v_A = \frac{1}{4}$	$v_A = \frac{1}{2}$	$v_A = \frac{3}{4}$			$v_A = \frac{1}{2}$
1	1	1	1	1	1	1	1
1	2	1·20	1·12	1·20	2	1	1/1·12
1	5	1·60	1·80	1·60	5	1	1/1·80
1	10	2·52	3·02	2·52	10	1	1/3·02
1	100	19·4	25·5	19·4	100	1	1/25·5
1	1000	187·5	250	187·5	1000	1	1/250

5. Discussion

The present lattice model demonstrates that considerable information can be obtained regarding special aspects of flow in heterogeneous media. However, as already noted, there is little experimental evidence concerning the extent to which regular lattice models can represent those heterogeneous media in which shape, size and arrangement of the particles of disperse phase is random. Randomization of the present model would lead to exchange of flux between neighbouring chains at points where one chain is more conducting than its neighbour. This effect tends to enhance flow but at the same time increases the tortuosity of flow paths compared with the regular model so that the overall result is not easily predictable.

In a more general version of the lattice model of Fig. 1, the parallelepipeds of phase A are of oblong rather than square cross section, of dimensions a and c, the corresponding dimensions for the unit cell being $a + b$ and $c + d$. This model again gives Eqn (18) and all its special cases, except that v_A is equal to $l_A ac/l(a + b)(c + d)$. Alternative regular lattices not having identical properties with the ones discussed so far are also possible. Two are shown in Figs 2(*a*) and 2(*b*), in which the chains of Fig. 1 are staggered in either of two ways, the case illustrated being for $kD_A \gg D_B$. When $kD_A \ll D_B$, the main feature of Fig. 2(*b*) is the greater tortuosity of flow paths and hence lesser permeability than arises with either Fig. 1 or Fig. 2(*a*).

It has recently been shown that for gas flow in microporous media, diffusion coefficients characteristic of steady and transient states of flow may differ appreciably (Barrer and

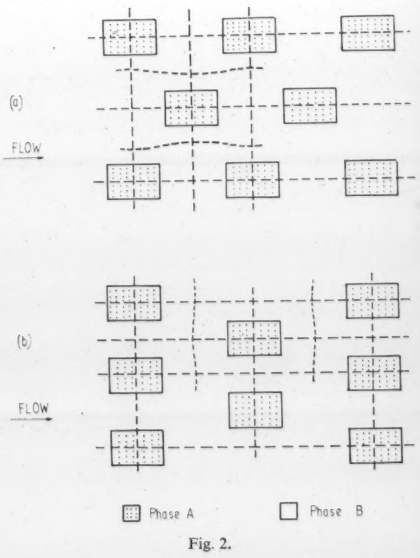


Fig. 2.

continuous layers normal to the direction of flow. Alternating layer structures also appear as particular cases of the lattices considered in this paper (§3.2.2. (i)). They have the same values of $\bar{D}$ in steady and transient states for all values of kD_A and D_B (Eqn (30)). Thus time lags are readily calculated for such media from the relevant equations.

Two approximate treatments of flow in cubic lattice models have been given previously in connection with the conduction of heat in porous solids (Topper 1955) and the magnetic permeability of iron powder compact cores for coils (*Wireless Engineer* 1933). The equation given by Topper, when there is negligible transfer of heat by radiation and convection, is the same as Eqn (34) of the present paper, although this author does not appear to have considered limitations in its validity and in fact used it with the condition $D_B \gg kD_A$, and irrespective of v_A. The formula given for iron powder magnetic cores, for which the condition of large v_A and kD_A/D_B is fulfilled, resembles our Eqn (35), but tends to give somewhat lower values for $\bar{D}$.

In the introduction it was noted that mass diffusion is paralleled in other physical processes. The equivalents of the diffusion flux J for these processes are respectively the heat flow J_{Th}, current flow I, electric displacement $\mathscr{D}$, and magnetic flux B. We note for the diffusion fluxes that

$$\frac{\bar{J}}{J_B} = \frac{\bar{D}(kv_A + v_B)}{D_B}; \quad \frac{J_B}{J_A} = \frac{D_B}{kD_A}$$

and, omitting the surface barrier term from the general Eqn (16), we rearrange this equation as

$$\frac{\bar{J}}{J_B} = \frac{h}{h_A(a+b)^2}\left\{\frac{a^2\beta_A}{\dfrac{J_B}{J_A} + \dfrac{h_B}{h_A}\dfrac{\beta_A}{\gamma\beta_B}} + \frac{b(2a+b)\beta'_A}{1 + \dfrac{h_B\beta'_A}{h_A\gamma'\beta'_B}}\right\}.$$

The relevant substitutions for $\bar{J}/J_B$ and J_A/J_B in the cases of the other physical processes are then obvious.

Acknowledgments

One of us (J.H.P.) wishes to acknowledge financial support through a post-doctoral fellowship from the Esso Petroleum Company Limited.

References

ASH, R., and BARRER, R. M., 1959, *Phil. Mag.*, **4**, 1197.
BARRER, R. M., 1940, *Trans Faraday Soc.*, **36**, 1235.
—— 1951, *Diffusion in and through Solids* (Cambridge University Press), p. 412.
BARRER, R. M., and GABOR, T., 1959, *Proc. Roy. Soc.* A, **251**, 353.
BÖTTCHER, C. J. F., 1945, *Rec. Trav. Chim. Pays Bas*, **64**, 67.
BROWN, W. F., Jr., 1955, *J. Chem. Phys.*, **23**, 1514.
BRUGGEMAN, D. A., 1935, *Ann. Phys., Lpz.*, **24**, 636.
EICHBAUM, B. R., 1959, *J. Electrochem. Soc.*, **106**, 804.
JAEGER, J. C., 1950, *Quart. J. Appl. Maths*, **8**, 187.
KLUTE, C. H., 1959, *J. Appl. Pol. Sci.*, **1**, 340.
MAXWELL, J. C., 1873, *Treatise on Electricity and Magnetism* (Oxford: Clarendon Press), 1st edn, Vol. 1 (3rd edn, 1904, Vol. 1, p. 440).
MEREDITH, R. E., and TOBIAS, C. W., 1960, *J. Appl. Phys.* **32**, 1271
NIESEL, W., 1952, *Ann. Phys., Lpz.*, **6**, 336.
—— 1953, *Ann. Phys., Lpz.*, **6**, 410.
POLEY, J. P., 1953, *Physica*, **19**, 298.
PRAGER, S., 1960, *J. Chem. Phys.*, **33**, 122.
RAYLEIGH, LORD, 1892, *Phil. Mag.*, **34**, 481.
RUNGE, I., 1925, *Z. Tech. Phys.*, **6**, 61.
SILLARS, R. W., 1937, *J. Instn Elect. Engrs*, **80**, 378.
TOPPER, L., 1955, *Ind. Engng Chem.*, **47**, 1377.
VAN BEEK, L. K., 1960, *Physica*, **26**, 66.
Wireless Engineer, 1933, **10**, 1.

Photoelectric processes in calcium oxide coated cathodes

by C. H. B. MEE,* B.A., M.Sc., Ph.D., Grad.Inst.P., and F. A. VICK,† O.B.E., Ph.D., A.M.I.E.E., F.Inst.P., Physics Department, University College of North Staffordshire, Keele, Staffs.

MS. received 17*th July* 1961, *in revised form* 7*th September* 1961

Abstract

The spectral sensitivities of photoelectric emission and photoconductivity in probe diodes containing cathodes coated with calcium oxide have been determined over the range of photon energies from 1·4 *to* 3·1 ev. *The two spectral sensitivities are similar in form, showing a rapid rise below about* 2·2 ev, *followed by a more gradual increase. Structure at about* 2·5–2·6 ev *is insensitive to the state of thermionic activity of the cathode, and is attributed to the presence of an impurity level at this depth below vacuum potential in the optical energy level scheme. Experiments involving studies of the recovery of photoelectric emission and photoconductivity after oxygen poisoning and of magnetoresistive effects in dark and illuminated conductivity suggest that the greater part of the photoconductivity current is carried by a process of photoelectric emission across the cathode pores. Thermal activation energies have been determined by measurements of the thermionic emission from, and the electrical conductivity of, the cathodes as functions of temperature.*

1. Introduction

METHODS based on the study of the spectral sensitivities of photoelectric emission and photoconductivity in barium oxide coated cathodes have been employed (Apker, Taft and Dickey 1951, DeVore and Dewdney 1951, Sakamoto 1954, Philipp 1957) in attempts to determine the energy band structure of the semiconducting oxide. The structure in the spectral sensitivity curves has been related to discrete energy levels originating from impurity levels in the oxide and to levels in the filled band. Early methods of determining the so-called photoelectric work function of the oxide suffered from a legacy of work on photoelectric emission from metals. Thus work functions determined by the method of complete photoelectric emission (Suhrmann and Fruehling 1938) have little theoretical significance. Spectral sensitivity results have been fitted to a Fowler plot for photoemission from metals (Mahlman 1949) but agreement with the theoretical curve is obtained over a limited range only. Work functions deduced from cut-off frequencies (Nisibori, Kawamura and Hirano 1941) depend on the intensity of illumination of the specimen, on its temperature and on the sensitivity of the instrument used to measure the photocurrent. Such methods often gave optical activation energies in reasonable agreement with thermal activation energies obtained from Richardson plots for thermionic emission and from measurements of electrical conductivity as a function of temperature, but it must be pointed out that this largely fortuitous agreement is no justification for ignoring the Franck–Condon effect. Unfortunately, there is insufficient experimental evidence available to make an estimate of the magnitude of the effect in this case.

A feature of one of the proposed optical energy level diagrams for barium oxide, due to Sakamoto (1954), is that the surface work function of the oxide is practically zero. This assumption is based on the experimental evidence of close similarities at corresponding values of the photon energy between the spectral sensitivity curves for photoelectric emission and for photoconductivity. Sakamoto has made the suggestion, but has not applied experimental tests, that the similarity may be due to the fact that photoconductivity in the porous cathode coating may be essentially a process of photoemission across the pores, instead of being a true volume or surface photoconductive effect. Such a hypothesis of pore photoconductivity has a parallel in the Loosjes–Vink theory (Loosjes and Vink 1949) of electrical conductivity in the oxide cathode at high temperatures, where the fundamental process of conductivity has been established to be one of free electron flight across the pores.

The present work was undertaken in an attempt to investigate the optical activation energies of calcium oxide which may lie in or near the visible region. Calcium oxide has received much less attention than the other alkaline earth oxides, and the thermal activation energies for calcium oxide coated cathodes have only recently been investigated (Hopkins and Vick 1958). A secondary aim of the work has been to confirm these values. It has also been possible to investigate the pore photoconductivity hypothesis of Sakamoto experimentally using methods adapted from some of those used to confirm the Loosjes–Vink hypothesis. These include an investigation of the recovery of photoelectric emission and photoconductivity after poisoning attacks by oxygen (the recovery of thermionic emission and electrical conductivity after poisoning has been investigated by Shepherd (1953) and by Higginson (1957)) and a study of the behaviour of photoconductivity currents in a magnetic field (magnetoresistive effects in dark conductivity have been studied by Forman (1954) and by Metson (1959)).

2. Apparatus and techniques

2.1. *Experimental tubes*

Photoelectric emission and photoconductivity currents have been measured in probe diodes containing a calcium oxide coated cathode. This type of tube, and its construction and processing, have been fully described by Hopkins and Vick (1958), and the details of preparation will not be repeated here. Basically, the cathode consists of a nickel tube coated over a length of about 2 cm with calcium oxide to a thickness of about 100 μ. In this coating is embedded a helical probe

* Now at the Physical Laboratory, University of Southampton.
† Now at the Atomic Energy Research Establishment, Harwell, Didcot, Berks.

of fine nickel wire, attached to insulated nickel wire supports. The cathode base is provided with a tungsten thermocouple lead and an insulated tungsten hairpin heater. The cylindrical nickel anode is pierced with a small window. In the standard type of probe diode (designated the PC series in this work) the glass frame bearing the electrode structure is mounted on a pinch and is sealed into an envelope so that the cathode lies along the axis of the envelope, but in certain applications in the present work a shorter cathode base and cathode coated length are employed and the diode assembly is attached to the pinch so that the cathode is perpendicular to the axis of the envelope. Tubes of this type make up the M series. Seventeen probe diodes (of both series) have been constructed during the course of the present work.

2.2. *Measurement of thermionic emission and electrical conductivity*

Measurements of the thermionic and electrical properties of the calcium oxide coating were carried out using the methods described in detail by Hopkins and Vick (1958). The small currents involved (down to 10^{-12} A) were measured with an electrometer designed around the Ferranti BDM 10 double-beam tetrode, and particular care was taken with the screening of the experimental tube. For most measurements the tube was placed in a heavy steel box, a small window permitting the illumination of the cathode for photoelectric measurements.

2.3. *Light sources and photoelectric measurements*

White-light illumination was provided from a 500 w tungsten filament lamp, mounted in a fan ventilated lamphouse. Light from the 500 w lamp was focused on the experimental cathode through the anode window.

Monochromatic illumination in the wavelength range 9000–4000 Å (photon energies of 1·4–3·1 ev) was obtained using a Hilger constant deviation spectrometer, Model D.186, as a monochromator in conjunction with the tungsten filament lamp. The experimental tube, in its screened box, was placed at the exit slit of the monochromator and its position adjusted until the cathode received maximum illumination through the anode window. Photocurrents were measured by applying potentials to either the anode or the probe and comparing electrometer readings with the shutter at the entrance slit closed and open. Several readings of the photocurrent were taken at each wavelength to allow for the effects of zero drift in the electrometer and to give a measure of the probable error in each determination. Photocurrents measured in this way must be divided by the relative intensity of illumination at the relevant wavelength to obtain values of the photocurrent per unit incident intensity. A calibration of relative intensity of illumination versus wavelength was performed using a photoemissive cell of known spectral sensitivity at the exit slit of the monochromator and determining the photocurrent over the spectral range. The intensity of illumination falling on the cathode of the experimental tube is not, however, identical with the intensity at the exit slit, due to absorption in the glass envelope of the tube and in the film of material evaporated on to it during the operation of the tube. An approximate correction for this absorption was carried out by taking the average of two sets of photocell readings: first with the photocell placed at the exit slit, and then with the envelope of the experimental tube interposed between the exit slit and the photocell. The relative intensity of radiation striking the experimental cathode at any wavelength is found, approximately, by dividing the average photocurrent by the sensitivity of the photocell at that wavelength.

2.4. *Oxygen poisoning experiments*

A number of the experimental diodes included a subsidiary alumina-insulated tungsten filament, coated with a paste of barium peroxide in ethyl acetate and collodion, mounted adjacent to the anode window. The filament was carefully outgassed during the processing of the cathode, and by passing a current through the outgassed filament and heating it to redness the peroxide could be broken down to give oxygen, barium oxide remaining on the filament.

Experiments to investigate the recovery of photocurrents after oxygen poisoning were carried out as follows. The cathode was illuminated with white light from the tungsten filament source and a value for the initial unpoisoned photocurrent (either photoelectric emission or photoconductivity) was obtained by applying potentials to the anode or to the probe. Oxygen was released by heating the poisoning filament for an arbitrary period of the order of 10 seconds, by which time the photocurrent would normally show a considerable reduction. The photocurrent was then monitored as a function of time after the current in the poisoning filament was cut off, and a recovery curve was traced.

2.5. *Conductivity in a magnetic field*

The experimental tube was set up with its cathode lying in the direction of the magnetic field between the pole-pieces of a Newport Type A Clarendon design electromagnet. The magnetic field was thus perpendicular to the approximately radial electric field between the probe and cathode base. The magnetic field as a function of the current in the coils of the electromagnet was determined for two separations of the pole-pieces, 6·5 cm (for tubes of the M series) and 21 cm (for the PC series), using a search coil and fluxmeter.

The effect of the magnetic field on dark conductivity currents could be investigated by determining the approximate conductivity in zero magnetic field by applying a potential of, say, 100 mv between probe and cathode base, reversing the polarity of the potential and observing the difference in electrometer readings; this procedure was repeated when a current was passed through the energizing coils of the electromagnet. The zero-field conductivity was redetermined after the energizing current had been cut off.

A similar procedure was followed in the investigation of the effect of the magnetic field on photoconductivity currents. The photoconductivity current for white-light illumination for a fixed probe potential was determined first in zero magnetic field and then in a magnetic field of the desired magnitude. It was not found possible to investigate the effect for photocurrents induced by monochromatic illumination, as the currents to be measured when the cathode is illuminated by the monochromator are less by a factor of about 10^4 than those produced by white-light illumination. The electrometer was found to be less stable in these experiments than in those where the tube could be adequately screened, and such small currents could not be measured accurately.

3. Experimental results

3.1. *Thermionic emission and electrical conductivity*

Thermal activation energies for the cathodes under investigation have been obtained from the gradients of conductivity

plots and of Richardson plots for thermionic emission, using the methods described by Hopkins and Vick (1958). Emission characteristics are obtained at a number of cathode temperatures T and the zero-field emission current i_0 is determined from the point of intersection of the retarding and accelerating field characteristics. The emission current density is sufficiently low for space-charge effects to be small, even at the highest temperatures employed (about 950° K), and saturation of the characteristics is good. A Richardson plot of $\log i_0/T^2$ against $1/T$ may be prepared and normally approximates to a straight line; the gradient of the line gives the Richardson work function ϕ_R.

The electrical conductance σ of the cathode coating at any temperature may be obtained by tracing a conductivity current–voltage characteristic for small potentials applied between the probe and the cathode base. The gradient of the characteristic at the origin is a measure of σ. Graphs of $\log \sigma$ against $1/T$ show the characteristic Loosjes–Vink pattern previously observed by Hopkins and Vick (1958) for calcium oxide coated cathodes, showing the existence of two temperature-dependent conductivity mechanisms acting in parallel. Hopkins and Vick used the symbols Q_1 and Q_2 to represent the activation energies of the high- and low-temperature processes respectively; $Q_{1\text{corr.}}$ represents the value of Q_1 corrected for the parallel low-temperature mechanism by the simple method described by Shepherd (1953).

The average Richardson work function for the most active states of eight cathodes prepared in the present work was $\bar{\phi}_R = 1{\cdot}6 \pm 0{\cdot}2$ ev. The average gradients of the conductivity plots, also for the most active states, corresponded to $\bar{Q}_1 = 1{\cdot}4 \pm 0{\cdot}2$ ev, $\bar{Q}_2 = 0{\cdot}8 \pm 0{\cdot}2$ ev and $\bar{Q}_{1\text{corr.}} = 1{\cdot}7 \pm 0{\cdot}2$ ev. These values should be compared with those quoted by Hopkins and Vick, where the average Richardson work function for five cathodes was $1{\cdot}7 \pm 0{\cdot}1$ ev, and the conductivity plots for these cathodes yielded values of $\bar{Q}_1 = 1{\cdot}1 \pm 0{\cdot}1$ ev, $\bar{Q}_2 = 0{\cdot}8 \pm 0{\cdot}1$ ev and $\bar{Q}_{1\text{corr.}} = 1{\cdot}3 \pm 0{\cdot}1$ ev. A further comparison may be made with conductivity results obtained with samples of calcium oxide contained between two flat nickel buttons, spring-loaded in contact (Mee 1961). The values obtained were $\bar{Q}_{1\text{corr.}} = 1{\cdot}2 \pm 0{\cdot}2$ ev and $\bar{Q}_2 = 0{\cdot}7 \pm 0{\cdot}1$ ev.

3.2. *Photoelectric emission*

A typical current–voltage characteristic for photoelectric emission caused by white-light illumination is shown as the continuous curve in Fig. 1. Photocurrents saturate in both directions; it is of interest to note that photocurrents are observed both from the cathode and from the anode, which

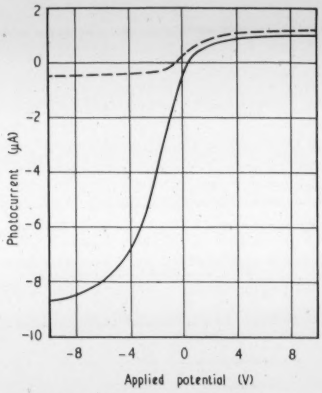


Fig. 1. Photoelectric emission and photoconductivity current–voltage characteristics for activated calcium oxide under white-light illumination. ——— photoelectric emission at 300° K; — — — photoconductivity at 300° K.

is illuminated by reflection from the cathode, and are of such a magnitude as to indicate that the photosensitivity of the anode must be considerably greater than that of the cathode. This photosensitivity is attributed to the presence of a thin film of material evaporated from the cathode during activation and operation. The photoelectric properties of such films have been the subject of a further investigation, the results of which will be reported in a subsequent paper.

Typical spectral sensitivity curves for photoemission from two cathodes are shown in Fig. 2, in which the photocurrents per unit incident intensity, i_{ph}, have been plotted against linear and logarithmic axes. The photocurrents have been expressed in arbitrary units, but in the typical cases shown the currents at an incident wavelength of 5600 Å (2·23 ev) for an incident power of about 10^{-6} w were about 4×10^{-12} A and 10^{-11} A for PC 5 and PC 9 respectively, corresponding to about 8×10^{-7} and 2×10^{-6} photoelectrons per quantum of incident light.

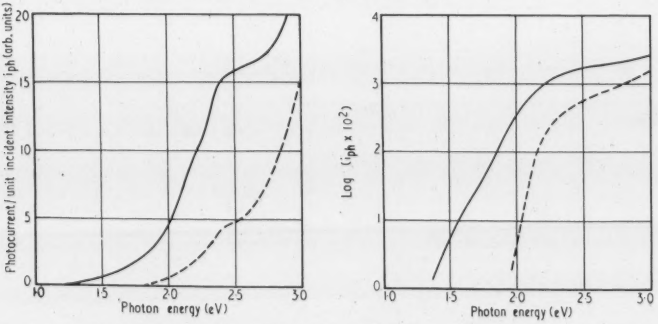


Fig. 2. Spectral sensitivity of photoelectric emission from calcium oxide at 290° K. ——— tube PC 5; — — — tube PC 9.

The curves obtained for the various cathode activation states generally show considerable similarity, with a fairly rapid increase in spectral sensitivity with increasing incident photon energy up to about 2·2 ev, where the increase becomes more gradual. There are some indications of structure in the spectral sensitivity curves between about 2·5 and 2·6 ev; this is more apparent in the curves drawn against linear than logarithmic scales. The 2·5–2·6 ev structure is observed in all states of thermionic activation, and is apparently insensitive to changes in the thermionic work function; on the other hand, it is generally found that cathodes with low thermionic work functions show an extended tail in the spectral sensitivity curves towards low photon energies. The Richardson work functions of the cathodes which yielded the curves in Fig. 2 were 1·0 ev (PC 5) and 2·4 ev (PC 9), and the spectral sensitivity curve for PC 5 extends considerably further towards low photon energies than does that for PC 9.

The temperature variation of white-light photoemission has been investigated by illuminating the cathode with white light from the tungsten filament lamp and determining the photocurrent as a function of the temperature of the cathode base. At temperatures above about 500° K it was necessary to correct the photocurrent for a thermionic emission component, and at temperatures above about 800° K the thermionic current became much larger than the photocurrent. A typical temperature variation curve is shown in Fig. 3.

3.3. *Photoconductivity*

Current–voltage characteristics for photoconductivity induced by white-light illumination also saturate in each direction. A typical curve is shown in Fig. 1. Potentials of up to 10 v in each direction have been applied between probe and cathode base to display this effect. In many cases it was found that the photoconductivity current in one direction was very much less than that in the other, and in these cases it was possible to determine the spectral sensitivity of photoconductivity for current flowing in one direction only. Experiments on cathodes in which it could be determined in both directions indicated that the spectral sensitivities of photoconductivity currents to and from the probe ('positive' and 'negative' photoconductivity currents) were the same, within experimental error. Fig. 4 illustrates the spectral sensitivity curves for photoconductivity in two cathodes. The form of the curves is very similar to that for photoelectric emission (Fig. 2), showing the same rapid increase at low photon energies, with just resolvable structure at 2·5–2·6 ev.

The temperature variation of white-light photoconductivity

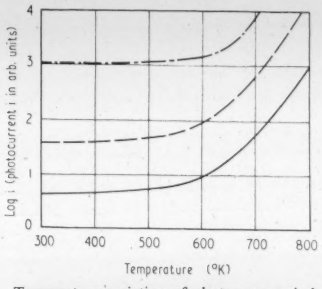


Fig. 3. Temperature variation of photocurrents induced by white-light illumination. Tube PC 5:
– · – · – · negative photoconductivity current,
— — — positive photoconductivity current,
———— external photoelectric emission (curve displaced downwards).

may be determined by tracing the illuminated and dark conductivity current–voltage characteristics at a number of temperatures. The current due entirely to illumination can be obtained by subtracting the dark conductivity current from the photoconductivity current at corresponding probe voltages. The temperature variations of the saturated positive and negative photoconductivity currents are similar (Fig. 3) and bear a marked resemblance to the temperature variation of external photoelectric emission.

3.4. *Oxygen poisoning experiments*

Figure 5 shows the results of an experiment in which the white-light photoelectric emission and photoconductivity currents were monitored simultaneously as a function of time after oxygen poisoning. The vertical axis represents the photocurrents expressed as percentages of the original unpoisoned photocurrents, and the time scale has been measured from the moment of switching off the current in the poisoning filament. It is observed that photoelectric emission and photoconductivity follow a similar recovery curve.

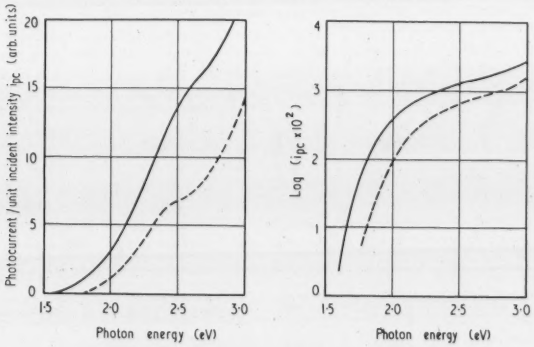


Fig. 4. Spectral sensitivity of photoconductivity in calcium oxide at 290° K.
———— tube PC 5; — — — tube PC 9.

It should, however, be pointed out that the form of the recovery curve is apparently a function of the partial pressure of oxygen, as curves of different form have been obtained for

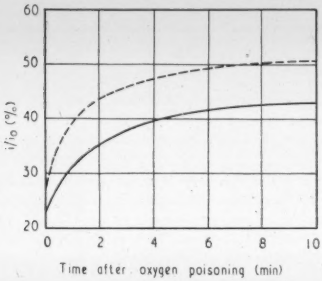


Fig. 5. Recovery of photocurrents after oxygen poisoning.
——— photoelectric emission at 290° K;
– – – – photoconductivity at 290° K (tube PC 7).

experiments in which the filament was flashed for varying lengths of time. In each case, however, the photoemission and photoconductivity currents recovered in a similar manner.

3.5. *Conductivity in a magnetic field*

Figure 6 shows the dependence of $f(\sigma) = (\sigma_0 - \sigma_H)/\sigma_0$ on H, where σ_0 is the conductivity in zero magnetic field and σ_H is the conductivity in magnetic field H, determined at a cathode temperature of 950° K (tube M 5; continuous curve in Fig. 6). The break point in the conductivity curve

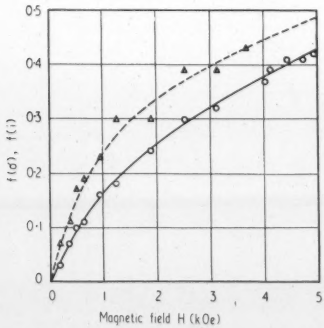


Fig. 6. Magnetoresistive effect in dark and illuminated conductivity. ——— dark conductivity at 950° K, vertical axis, $f(\sigma)$; – – – – illuminated conductivity at 290° K, vertical axis, $f(i)$ (tube M 5).

for this cathode occurred at 780° K; thus, on the Loosjes–Vink theory, the electrical conductivity of the cathode at temperatures above 780° K is largely due to electron emission across the pores of the cathode coating, and the pronounced magnetoresistive effect displayed at 950° K is due to the superposition of a transverse drift velocity on the electrons travelling through the cathode matrix. Figure 7 shows the temperature variation of $f(\sigma)$ determined at a fixed value of H of 1000 oersteds, together with the conventional conductivity plot (tube PC 5). At temperatures below the break in the conductivity plot, $f(\sigma)$ becomes very small.

A magnetoresistive effect is also observed in photoconductivity at room temperature. In Fig. 6 $f(i) = (i_0 - i_H)/i_0$ has been plotted against H as a broken line, where

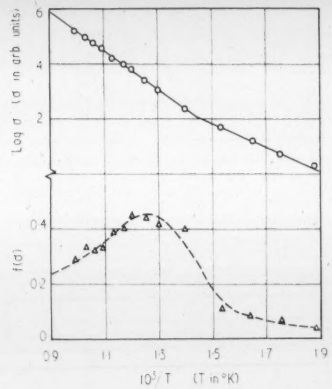


Fig. 7. Temperature variation of the dark magnetoresistive effect (– – – –) and of electrical conductivity (———) (tube PC 5).

i_0 and i_H are the photoconductivity currents in zero magnetic field and in magnetic field H, respectively (tube M 5). There is considerable scatter of the points about the curve drawn, due to difficulties in measuring small currents from the unscreened tube, but the form of the curve is sufficiently clear to show its similarity to that for dark conductivity in the same cathode in the pore conduction region (the continuous curve in Fig. 6). A similar curve of $f(i)$ plotted against H is obtained whether photocurrents from the cathode base to the probe, or from the probe to the cathode base, are measured.

The possibility exists in all the experiments relating to photoconductivity currents that what has been considered to be a photoconductivity current to and from the probe is a photoemission current to and from the probe supports, which are external to the cathode surface. Among the tubes constructed was one in which particular precautions were taken against the illumination of the probe supports, either directly or by reflection from the cathode and anode. In this tube photocurrents were observed both to and from the probe and the magnetoresistive effect in illuminated conductivity was obtained for both positive and negative photocurrents. If photoemission from the probe supports was an important factor, negative photocurrents would not have been observed (or would have been very much diminished) in this tube. The fact that they were observed suggests that any component of the negative photoconductivity current due to photoemission from the probe supports must be very small compared with that originating in the region of the cathode near the embedded probe.

4. Discussion

4.1. *Photoelectric emission and photoconductivity*

Spectral sensitivity curves of photoelectric emission and photoconductivity show a strong similarity in form. Each displays a rapid rise in photosensitivity below about 2·2 ev

and a more gradual increase above this figure, and in addition some structure is obtained at 2·5–2·6 ev in each set of curves. A closer examination of the logarithmic spectral sensitivity curves for each cathode shows that they can be approximately superimposed by applying vertical and horizontal translations, the horizontal translation in all cases being small (of the order of 0·1 ev). The close similarity of photoelectric emission and photoconductivity spectral sensitivity results recalls the interpretations which have been put forward by Sakamoto (1954) as a result of his investigations on barium oxide cathodes. The small horizontal translation required to bring photoelectric emission and photoconductivity curves into approximate coincidence may be interpreted on conventional semiconductor theory by supposing the optical external work function to be very small, of the order of 0·1 ev, or alternatively by assuming that the photoconductivity currents observed in the porous cathode are in fact photoemissive currents across the cathode pores. The latter interpretation is strongly favoured by the experimental evidence described in § 3.4 and § 3.5, and is discussed in further detail in § 4.2.

Considerable caution is required in deducing optical activation energies from the spectral sensitivity curves. Cut-off photon energies can be obtained from the sensitivity curves drawn against linear axes, but cannot yield activation energies of any significance. The curves drawn against logarithmic axes are nearly linear at low photon energies and suggest that if a more intense source of illumination or a more sensitive electrometer had been used a lower cut-off energy would have been obtained. Arbitrary thresholds obtained from these cut-off energies show a rough correlation with thermal activation energies obtained from Richardson and conductivity plots for corresponding states, low values of the photoelectric thresholds being associated with low thermal activation energies. This can merely be taken as an indication that thermionic activity and photosensitivity are correspondingly greater in the more active states of the cathode.

It is suggested that the structure observed in the experimental curves for photoelectric emission at about 2·6 ev may be due to the existence of an impurity level at this depth below vacuum potential. This level cannot, however, be the one associated with thermionic emission and electrical conductivity, as the structure is observed at approximately the same photon energy in all states of cathode activation over a wide range of values of the Richardson work function.

It should be pointed out that it is unlikely that any structure observed in the present work should have arisen from intrinsic transitions from the filled band. The calcium oxide cathode is normally white, showing that there is no considerable change in optical absorption in the visible region (1·6–3·0 ev) to correspond with the intrinsic absorption edge. It may, however, be mentioned that on storing activated calcium oxide cathodes a blue coloration frequently occurs: this has been reported previously by Hopkins and Vick (1958).

4.2. *Pore photoconductivity*

Shepherd (1953) and Higginson (1957), who studied the recovery of thermionic emission and electrical conductivity in oxide coated cathodes after oxygen poisoning, found that at temperatures above the break in the conductivity plot emission and conductivity recovered in the same characteristic manner; at temperatures below the break the recovery curves were quite different. This information has been used to strengthen the Loosjes–Vink theory of pore conduction at high temperatures. In the present work, the similarity of the recovery curves of photoelectric emission and photoconductivity after oxygen poisoning at room temperature suggests that at this temperature the photoconductivity current originates in a photoemissive process. That the photoemissive nature of photoconductivity may extend over a considerable temperature range is suggested by the similarity in form between the temperature variations of photoelectric emission and photoconductivity (Fig. 3).

Forman (1954) and Metson (1959) have noted that at temperatures above the break in the conductivity plot for barium oxide and mixed oxide cathodes a marked magnetoresistive effect is observed. At temperatures below the break no effect is measurable. This behaviour, also, has been intrepreted in favour of the Loosjes–Vink theory. The present work shows that a similar magnetoresistive effect is observed in dark conductivity in calcium oxide coated cathodes, and helps to confirm that calcium oxide cathodes are pore conductors at high temperatures. The close similarity between the graph of $f(\sigma)$ against H for dark conductivity in the pore conduction range and that of $f(i)$ against H for photoconductivity at room temperature suggests that photoconductivity is a pore phenomenon at room temperature.

The experimental results thus lead to the conclusion that photoconductivity in calcium oxide coated cathodes at room temperature is effectively a process of photoelectric emission across the pores of the cathode. Thus a study of the spectral sensitivity curves for photoelectric emission and photoconductivity can only yield information referring to the position of energy levels relative to vacuum potential, and it is not possible to deduce the optical surface work function from the results of the present work. Slight differences between the spectral sensitivity curves for photoelectric emission and photoconductivity in the same cathode may be ascribed to differences in the emitting surfaces. The photoelectric emission results are bound to be affected by the cathode layers near the surface, outside the probe, where a proportion of the incident light is absorbed, but this region of the cathode can have no effect on the photoconductivity results.

4.3. *Thermionic emission and electrical conductivity*

The mean Richardson work function for thermionic emission from activated calcium oxide cathodes and the value of $\bar{Q}_2$, the mean activation energy of the low-temperature conduction process, given in § 3.1, are in agreement, within experimental error, with the values reported by Hopkins and Vick (1958) and by Mee (1961). The mean activation energy for the high-temperature conduction mechanism is, however, rather greater than those quoted in the earlier papers. In the present case the mean corrected high-temperature activation energy is equal, within experimental error, to the mean Richardson work function, and in agreement with the predictions of the Loosjes–Vink theory. If χ is the surface work function of calcium oxide and ΔE the depth of the impurity level associated with thermionic emission and electrical conductivity below the conduction band, the Richardson work function ϕ_R is approximately equal to $\chi + \Delta E/2$. The present results suggest the existence of an impurity level at a depth of about 1·6 ev below the bottom of the conduction band and 2·4 ev below vacuum potential, in the most active state of the cathode; the surface work function χ, obtained by direct subtraction of $\bar{Q}_2$ from $\bar{\phi}_R$, is about 0·8 ev. The relationship between thermal and optical activation energies cannot be deduced from the present results.

5. Conclusions

The porous structure of the calcium oxide coated cathode has been shown to play an important part in the photoelectric as well as in the electrical properties of the cathode. Comparative investigations suggest that the process of photoconductivity at room temperature is very similar to that of dark conductivity at higher temperatures, where the greater portion of the conductivity current is carried by a process of free electron flight across the pores of the coating.

Acknowledgments

The authors wish to express their thanks to Professor D. J. E. Ingram, who provided laboratory facilities during 1960, to Dr. N. A. Surplice, for helpful discussion during the course of the work, and to their colleagues and the technical staff of the Department. Siemens-Edison-Swan Ltd. made a generous gift of valve stems and envelopes.

References

APKER, E., TAFT, E., and DICKEY, J., 1951, *Phys. Rev.*, **84**, 508.
DEVORE, H. B., and DEWDNEY, J. W., 1951, *Phys. Rev.*, **83**, 805.
FORMAN, R., 1954, *Phys. Rev.*, **96**, 1479.
HIGGINSON, G. S., 1957, *Brit. J. Appl. Phys.*, **8**, 148.
HOPKINS, B. J., and VICK, F. A., 1958, *Brit. J. Appl. Phys.*, **9**, 257.
LOOSJES, R., and VINK, H. J., 1949, *Philips Res. Rep.*, **4**, 449.
MAHLMAN, G. W., 1949, *J. Appl. Phys.*, **20**, 197.
MEE, C. H. B., 1961, *Nature*, **190**, 1093.
METSON, G. H., 1959, *Proc. Inst. Elect. Engrs*, **106**C, 55.
NISIBORI, E., KAWAMURA, H., and HIRANO, K., 1941, *Proc. Phys. Math. Soc. Japan*, **23**, 37.
PHILIPP, H. R., 1957, *Phys. Rev.*, **107**, 687.
SAKAMOTO, M., 1954, *Vide*, **9**, 109.
SHEPHERD, A. A., 1953, *Brit. J. Appl. Phys.*, **4**, 70.
SUHRMANN, R., and FRUEHLING, G., 1938, *Naturwiss.*, **26**, 108.

Notes and comments

Vapour Pressure Tables of Non-associating Substances

Dampfdrucktabellen nichtassoziierender Stoffe (*Vapour Pressure Tables of Non-associating Substances*), edited by E. Oehley, has just been published as a new addition to the Dechema-Erfahrungsaustausch series.

The booklet provides the possibility of readily obtaining approximate information on boiling temperatures under a wide range of pressures, provided that the boiling point at 760 torr is known. It may be obtained from the Dechema Deutsche Gesellschaft für chemisches Apparatewesen E.V., Frankfurt (Main) 7, Postfach 7746. The price is DM 16.80 plus postage.

Frontier

We have received a copy of *Frontier* which is published by the Armour Research Foundation of Illinois Institute of Technology. The Armour Research Foundation celebrates its 25th Anniversary this year and *Frontier* has been redesigned and expanded partly as celebration and partly to enable it to cover the greatly increased research which the Foundation now carries out.

Frontier is published quarterly by the Armour Research Foundation of Illinois Institute of Technology, Technology Center, 10 W. 35th Street, Chicago 16, Illinois.

Nuclear Science Abstracts

Nuclear Science Abstracts is an abstracting and indexing service devoted to the literature of nuclear science and technology. It is published bi-monthly by the United States Atomic Energy Commission. It is available on a subscription basis from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. or on an exchange basis to universities, research institutions and publishers of scientific information from the Office of Technical Information Extension, U.S. Atomic Energy Commission, P.O. Box 62, Oak Ridge, Tennessee.

Microwave appraisal of maser crystals

by J. S. THORP, B.Sc., Ph.D., A.Inst.P.,* J. H. PACE and D. F. SAMPSON, Royal Radar Establishment, Malvern, Worcs.

MS. received 7th July 1961, in revised form 3rd August 1961

Abstract

The potentialities of microwave spectrometer techniques in assessing the purity and crystalline quality of synthetic maser crystals are illustrated by a comparison of the paramagnetic properties of ruby samples grown from the vapour phase and from powder by flame fusion. Typical spectra presented are discussed in relation to complementary chemical and spectrographic analysis, and it is shown the information obtained is of direct relevance to the accurate control of methods of crystal growth.

1. Introduction

THE increasing interest in low-noise maser amplifiers (Bloembergen 1961), both from the fundamental viewpoint and from that of developing complete microwave receiving systems incorporating their use, has emphasized the importance of the purity and perfection of the maser crystal and of the need for batch-to-batch reproducibility in its paramagnetic properties. In addition to the overriding conditions that the maser crystals should be capable of growth in single crystal form, possess the correct energy level scheme for the frequencies at which it is to be operated and have a reasonable relaxation time, a number of other properties are desirable if optimum maser performance is to be achieved. In the first place, the concentration of the required paramagnetic ion should be accurately controlled, since this determines the absorption line width and hence influences the bandwidth obtainable in any maser device; it can also affect the spin lattice relaxation time and may alter the pumping power required for maser operation at a given temperature. Secondly, the crystal should be free from unwanted paramagnetic impurities whose absorption lines coincide with or are very close to those of the required paramagnetic ion. If such impurities were present it would be possible for pump power to be wasted by saturating an unwanted transition, or for the relaxation time of the required transition to be seriously reduced by spin–spin coupling to an impurity transition having a very short relaxation time.

Microwave spectrometer techniques are particularly suitable as a means for studying these features, because measurements can be made under conditions similar to those experienced in maser operation. They are also useful in supplying more detailed information on the effects of crystalline perfection than can readily be obtained from visual inspection or simple polarized light and x-ray techniques. The measurements quoted in this paper refer specifically to synthetic ruby (Cr^{3+} in Al_2O_3) which, especially at wavelengths above 3 cm, is still widely used as a maser material, both on account of its relatively long relaxation time (Kikuchi *et al.* 1959, Pace *et al.* 1960) and because of its good mechanical and chemical stability. The techniques discussed in the following sections are, however, applicable to any solid maser material, and preliminary investigations are already in progress on chromium and iron doped rutile, which appear to be useful materials for millimetre wavelength masers (Foner *et al.* 1961).

* Now at Department of Applied Physics, Science Laboratories, Durham.

2. Microwave spectrometer appraisal

Technique

Since microwave spectrometer techniques are fairly well known, only the salient features of the method adopted will be given. Measurements were, for convenience, made with an 8 mm spectrometer which has previously been described (Pace *et al.* 1960). Within the fairly wide limits imposed by the ability to see transitions with a material of given zero field splitting the frequency of measurement is of secondary importance. On the other hand, the temperature should be either that at which the material is to exhibit maser action, or preferably lower, since if impurities whose transitions have very short relaxation times are present they will remain undetected at high temperatures, e.g. 77° K or room temperature, as their absorption lines may not be visible. In the present survey all measurements were made at 1·4° K. To simplify the interpretation of the spectra and to enable comparison to be made with some previous experiments on adiabatic rapid passage (Thorp *et al.* 1961) an orientation of $\theta = 90°$ was used throughout. The magnetic field values at which transitions occurred and their relative peak intensities were recorded; observations of the line shape were also taken.

Samples investigated

Details of the method of growth and nominal concentration of the samples investigated in the present survey are given in Table 1.

It should perhaps be pointed out that all the crystals examined were not in the fully annealed state, and it must also be borne in mind that the samples on which results are quoted were chosen to illustrate effects which may arise during the development of growth techniques, and do not necessarily represent the quality of crystal which may ultimately be obtainable.

Results

Typical results obtained on some of the samples listed above are shown in Figs 1 and 2, in which 0 dB corresponds to the noise level. In these the field values at which absorption lines occurred are given, together with the relative peak intensities. Where possible the lines were identified by comparison with the known spectra of chromium (Geusic

Table 1. *Details of growth method used in the samples investigated*

Sample	*Material*	*Remarks*
P 1	Ruby	Nominal 0·1% Cr ruby. Grown from powder by a flame fusion method.
P 2	Undoped Al_2O_3	Starting material for P 1. Grown by same method.
V 1	Ruby	Nominal 0·1% Cr ruby. Grown from vapour phase; Method I.
V 2	Undoped Al_2O_3	Starting material for V 1. Grown by same method.
V 3	Ruby	Nominal 0·1% Cr ruby. Grown from vapour phase; Method II.
V 4	Undoped Al_2O_3	Starting material for V 3. Grown by same method.

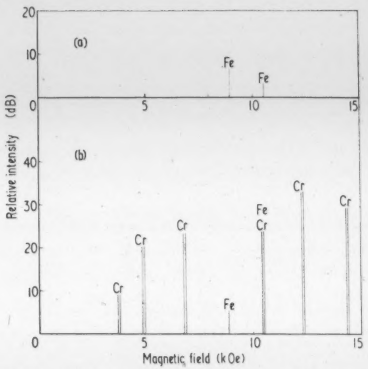


Fig. 2. Spectra of samples grown from vapour phase.
(*a*) Undoped aluminium oxide, V2; (*b*) Ruby, V1. Orientation $\theta \simeq 90°$, 34·6 Gc/s.

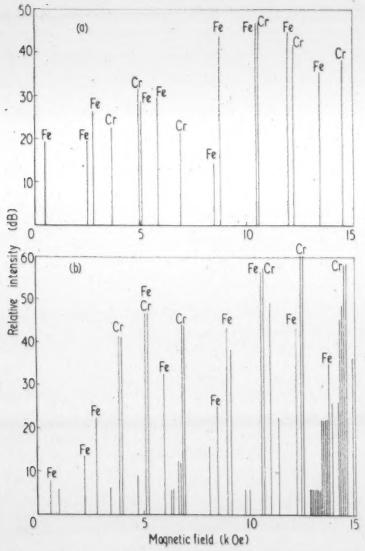


Fig. 1. Spectra of samples grown by flame fusion from powder.
(*a*) Undoped aluminium oxide, P2; (*b*) Ruby, P1. Orientation $\theta \simeq 90°$, 34·6 Gc/s.

1956, Howarth 1958, unpublished M.O.A. report) and iron (Bogle and Symmons 1959, Howarth 1960, unpublished M.O.A. report) in aluminium oxide; the iron lines could be uniquely identified because there are two non-equivalent Fe^{3+} ions in the unit cell and rotation of the magnetic field separated each line into a doublet, except at the symmetrical orientations giving coincidences. Samples V3 and V4 gave spectra almost identical to V1 and V2 respectively. The most striking feature is the comparative freedom from iron contamination in the crystals grown from the vapour phase, since both samples V1 and V3 were found to contain very little iron and no other detectable impurity. In these samples (V1 and V3) the maximum intensity of an iron transition is some 600 times (28 dB) less than that of the main chromium transitions. The sample P1, however, grown by flame fusion from powder, contains iron in considerably greater quantities and, in fact, shows Fe^{3+} transitions whose intensities are almost equal to those of some of the chromium lines; in addition, this sample gave numerous additional lines, some of significant intensity, which have not been identified. It is interesting to note the close correlation in the intensities of the Fe^{3+} spectra obtained from corresponding 'pure' and doped samples.

Observations of the line width and shape, which may be made either at low or room temperatures, give information regarding the crystalline perfection of the sample. It will be noticed from Figs 1 and 2 that, in some samples, chromium transitions, which should have been single lines since all the Cr^{3+} sites are identical, were in fact wholly or partially resolved doublets whose width was approximately 30 oersteds; these results showed that some form of twinning or macroscopic mosaic structure was present. Although this is not important in the appraisal of chemical purity, it would be very detrimental from the point of view of, for example, precise spectroscopic measurements, the attainment of optimum microwave maser performance and the homogeniety of refractive index in an optical maser sample.

3. Complementary analytical techniques

The spectra given in Figs 1 and 2 show that microwave spectrometer techniques form a sensitive means of detecting paramagnetic impurities. The method of identifying their nature by comparison is, however, limited to those ions whose paramagnetic properties in the host lattice are known. Hence a major requirement is for a complementary method of determining the nature of impurities which may be present; this method should also enable concentrations to be determined quantitatively, since it is not easy to make this measurement accurately by microwave methods. In this respect

chemical and optical spectrographic techniques offer the most promising approach, with polarographic and x-ray fluorescence as possible alternatives suitable for particular impurities. Work along these lines has been undertaken by the Chemical Inspectorate, Woolwich, with particular reference to the accurate determination of chromium content in a small sample (thus enabling a measure of the uniformity of doping of the crystal to be obtained) and to the determination of trace quantities of iron in the presence of chromium. Details of these techniques will be reported in separate publications. The analytical results for the present series of samples are given in Table 2.

Table 2. *Chromium and iron content of samples estimated by chemical and optical spectrographic techniques*

Sample	Chromium concentration		Iron concentration
	Chemical (atomic %)	Spectrographic (atomic %)	chemical (atomic %)
P1 ruby	0·038	0·033	0·035
P2 undoped Al_2O_3	<0·001	<0·005	0·039
V1 ruby	0·023	0·025	0·004
V2 undoped Al_2O_3	<0·001	<0·005	0·004
V3 ruby	0·011	0·008	0·004
V4 undoped Al_2O_3	<0·001	<0·005	0·004

In Table 2 the chromium and iron concentrations are expressed as atomic percentages, i.e. the number of Cr or Fe atoms per 100 molecules of Al_2O_3. These were derived from the weight percentages by multiplication by half the molecular weight of Al_2O_3 and by division by the atomic weights of chromium and iron respectively. It may be noted that the chromium concentrations in all the ruby samples are considerably smaller than the nominal concentrations, based on the proportions of additive to host in the starting materials, quoted in Table 1. It also appears that samples of the same nominal concentration, grown by different methods, yield true chromium concentrations which may differ by a factor of about three times. The difference between the nominal and true concentration arises from the difference in volatility of the host lattice material and the added paramagnetic material at the growing temperature, and the factor of three indicates the extent to which this may be influenced by various conditions of growth.

4. Discussion

The spectra of Figs 1 and 2 were obtained primarily to detect and identify impurities. Thus although the samples were of approximately the same volume and measurements were taken under similar conditions of coupling and receiver gain, further information regarding individual line widths and transition probabilities is necessary to enable detailed quantitative correlation with the analytical results to be made. The peak intensities observed can, however, be used to make some approximate comparisons. For example, in the crystals grown from the vapour phase, the iron content of the undoped Al_2O_3 (V2, Fe 0·004%) gave a maximum peak intensity about 5 dB above noise; this remained about the same in the corresponding ruby (V1, Fe 0·004%, Cr 0·024%) giving an intensity difference of about 28 dB between the strongest chromium and iron transitions. In the crystals grown from powder, the intensities of the corresponding iron transitions in both the undoped Al_2O_3 (P2, Fe 0·039%) and ruby (P1, Fe 0·039%, Cr 0·035%) were about 43 dB above noise. Both techniques thus indicated that the crystals grown from the vapour phase contained less iron than those grown from powder, and that the iron content of the rubies was substantially independent of the addition of chromium; they could not, however, differentiate between contamination which arose from impurities present in the starting materials or entered during growth. The difference in peak intensity of about 38 dB between corresponding iron transitions in P2 and V2 was considerably greater than expected from the ratio of their iron concentrations, and this suggested that there might be a significant difference between the amount of iron estimated analytically and that entering the lattice as Fe^{3+}. The paramagnetic spectra also showed variations in the relative intensities of transitions due to the same ion, which were more pronounced for Fe^{3+} than Cr^{3+}. It can be seen, however, that the information which can be derived from the paramagnetic spectra is directly relevant to crystal growth procedures, and that systematic study of samples grown under known conditions would be valuable in establishing growth techniques for very high purity crystals.

5. Conclusions

The use of a microwave spectrometer as a tool for the assessment of ruby maser crystals has been demonstrated. The chief advantage of the method was that the crystal could be examined under conditions which corresponded closely to those under which it would be required to operate in a maser. Comparison of ruby grown from the vapour phase with that grown by flame fusion from powder showed that, at the present stage of development of the growing techniques, the former was very much more free from impurities. When coupled with complementary chemical analytical techniques a powerful method of assessing material was obtained.

Acknowledgments

We are indebted to R. A. Mostyn and E. M. Dodson, The Chemical Inspectorate, Woolwich, for supplying the results given in Table 2. We also wish to acknowledge the ready co-operation of the Hirst Research Centre of the General Electric Co. Ltd., and the Thermal Syndicate Ltd. in supplying crystals on which measurements were made.

This paper is published by permission of the Controller, H.M. Stationery Office.

References

BLOEMBERGEN, N., 1961, *Progress in Low Temperature Physics*, **3**, 396.

BOGLE, G., and SYMMONS, H. F., 1959, *Proc. Phys. Soc.*, **73**, 531.

FONER, S., MOMO, L. R., THAXTER, J. B., HELLER, G. S., and WHITE, R. M., 1961, *Conference on Quantum Electronics, Berkeley* (New York: Columbia University Press).

GEUSIC, J. E., 1956, *Phys. Rev.*, **102**, 1252.

KIKUCHI, C., LAMBE, J., MAKHOV, G., and TERHUNE, R. W., 1959, *J. Appl. Phys.*, **30**, 1061.

PACE, J. H., SAMPSON, D. F., and THORP, J. S., 1960, *Proc. Phys. Soc.*, **76**, 697.

THORP, J. S., PACE, J. H., and SAMPSON, D. F., 1961, *J. Electronics and Control*, **10**, 13.

Determination of the emissivity, for total radiation, of small diameter platinum–10% rhodium wires in the temperature range 600–1450° C

by D. BRADLEY, B.Sc., Ph.D., A.M.I.Mech.E., and A. G. ENTWISTLE, B.Sc., Mechanical Engineering Department, University of Leeds, Leeds 2

MS. received 8th August 1961

Abstract

The method of measurement of the emissivities is described. Two different lengths of wire, of the same diameter, were electrically heated in vacuo. *By this procedure the end cooling correction could be eliminated experimentally. From the readings of wire resistance and current the net rate of radiant energy loss from the wire to the surroundings, and hence the wire emissivity, could be found.*

The values of emissivity are compared with those for platinum wires and also with values calculated from wire resistivity on the basis of electromagnetic theory.

1. Introduction

ALTHOUGH data exist on the emissivity of platinum wire at high temperatures, there have been few data available for platinum–10% rhodium, despite the use of this alloy in thermocouples and sometimes in resistance thermometers. For this reason the work was carried out.

2. Principles of the method

If a fine wire, mounted between larger diameter main leads is electrically heated *in vacuo* then, in the steady state, the electrical power supplied to the wire is equal to the sum of the net rate of radiative energy loss from the wire and the rate of conductive energy loss from the wire to the main leads. As the wire length is increased the rate of conductive energy loss becomes a smaller proportion of the power supplied. From measurements of the electrical power consumption it is possible to obtain the net rate of radiative energy loss and hence values of emissivity of the wire.

Davisson and Weeks (1924) used this method to find the emissivity of pure platinum wire of approximately 0·006 in. diameter. They employed two separate wires of different lengths and by means of an extrapolation of their experimental results, together with a theoretical analysis of the conduction loss, they allowed for the cooling effect of the main leads and also of the separate potential leads. They obtained values of emissivity at wire temperatures between 300° K and 1500° K.

A variation of this method was employed in the work here described. If the fine wire is surrounded by a container at a uniform temperature T_b then the wire is effectively exposed to black radiation from a body at this temperature and some of it will be absorbed. The energy equation for such an electrically heated wire, mounted between main leads *in vacuo*, with black body surroundings, is

$$\frac{\partial(k\partial T/\partial x)}{\partial x} + \frac{4i^2r}{J\pi d^2} - \frac{4\sigma}{d}(\epsilon T^4 - \alpha T_b^4) = 0 \qquad (1)$$

where T = temperature (° K) at distance x from one end of the wire, k = wire thermal conductivity, i = electric current through wire, r = wire electrical resistance per unit length at temperature T, J = mechanical equivalent of heat, d = diameter of wire, σ = Stefan–Boltzmann constant, ϵ = wire emissivity at temperature T, α = wire absorptivity at temperature T, to black body radiation from container walls.

Solutions of Eqn (1) for the boundary conditions $T = 300°$ K, the temperature of the mains leads, at $x = 0$ and $\partial\left(\frac{k\partial T}{\partial x}\right)/\partial x = 0$ at an infinite distance from the end of a wire of semi-infinite length were obtained (see § 5). For any required value of i, the temperature distribution along the length of the wire could be found and two of these distributions for semi-infinite platinum–10% rhodium of the diameter used in the experiments, are shown in Fig. 1. As both main leads were at the same temperature, the actual distribution for a particular wire was symmetrical about the mid-point. The mid-point of the shorter wire used is shown by the line AA in Fig. 1. The cooling effect of the main leads is clearly shown.

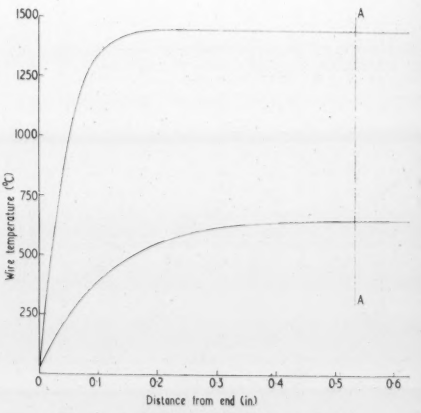


Fig. 1. Temperature distribution along wire.

Provided that the total length of the wire was sufficient to give a central region where the temperature was sensibly constant, any increase in wire length, other parameters remaining unchanged in value, resulted in an equal increase

in the length of the central region. The temperature distribution at the ends was the same in both cases. It was as if an extra length of wire at the constant central region temperature T_c had been added to the original wire. This was the basis of the experimental method.

Two wires, alike in all other respects but with a difference in length, X, were heated *in vacuo* and readings taken of the wire total resistance and the corresponding current. Provided that the length of the short wire was adequate then, for a given current, the two wires attained the same central temperature T_c. Let R_l and R_s be the resistance of the long and the short wires respectively, for the same wire current. As the 'extra' length of wire X may be regarded as being at the central region temperature T_c it follows that the wire electrical resistance per unit length r_c, at this temperature, is given by

$$r_c = \frac{R_l - R_s}{X}. \tag{2}$$

Over this length the first term in Eqn (1) is zero and the wire emissivity is given by

$$\epsilon = \frac{i^2 r_c}{J\pi d\sigma T_c^4} + \alpha\left(\frac{T_b}{T_c}\right)^4. \tag{3}$$

T_c in this equation was obtained from the experimentally derived values of r_c and the resistance–temperature relationship for the wire. The value of α was obtained theoretically and this is discussed in § 4. The term in α becomes less important as T_c is increased.

3. Experimental details

The platinum–10% rhodium wires had a mean diameter of 0·00057 in. The material was supplied on reels, in 50 ft lengths, by Messrs. Johnson Matthey of Hatton Garden.

The straight length of wire to be heated *in vacuo* was electrically welded between platinum–10% rhodium main leads of approximately 0·030 in. diameter. The main leads passed through a Perspex flange, the seal around the leads being Pizein wax. In the work described, the two lengths of wire were 1·425 in. and 1·064 in.

Wire diameters at room temperature were measured to an accuracy of ±0·00001 in. for each reading, using a Baker microscope, with a magnification of 400. Readings of diameter were taken at intervals of 0·1 in. along the length of the wire. Wire lengths were measured to an accuracy of ±0·001 in. using a Cooke, Troughton and Simms toolroom microscope, with a magnification of 50.

When calculating emissivities from Eqn (3), allowance was made for the expansion of the wires due to increase of temperature. Room temperature data on coefficients of expansion (*International Critical Tables* 1927) was extrapolated to higher temperatures, being guided in this by the known variation of the coefficients for platinum with temperature (*International Critical Tables* 1927).

The wire to be heated, welded to the main leads which passed through the Perspex flange, was mounted straight and horizontal in a steel container. This consisted of a 3½ in. high vertical cylinder of 5 in. internal diameter. A horizontal boss in the side of the cylinder with an opening of 1·125 in. internal diameter, allowed the introduction of the main leads and wire, which were then in a horizontal position. The Perspex flange, now vertical, with a brass backing ring on the outside was fastened on to the boss face with six 2 B.A. studs. An O-ring in a groove in the boss face effectively sealed the joint.

An Edwards rotary and oil diffusion pump unit evacuated the cylinder, which was fastened to the horizontal flange above the baffle valve. The pressure was read on a cold cathode ionization gauge mounted on top of the cylinder. During the heating of the wire the pressure was always less than 10^{-4} mm Hg and the wire could be observed through the Perspex flange. The temperature of the cylinder walls varied between 23° C and 25° C during the experiments.

Having obtained a sufficiently low pressure, the wire was electrically heated by direct current and the values of wire resistance R_w together with values of wire current i were obtained using the circuit shown in Fig. 2. The wire current

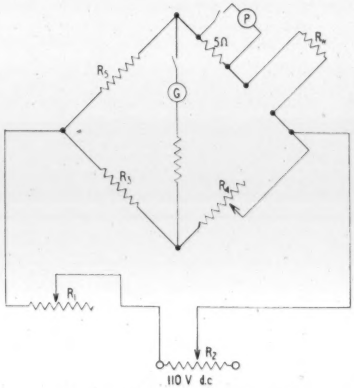


Fig. 2. Heating and measuring circuit.

could be adjusted by means of the variable resistances R_1 and R_2 whilst R_3 and R_5 were the known ratio resistances in a Wheatstone bridge. The wire resistance R_w was measured by noting the value of R_4 at balance of the bridge, indicated by no deflection of the Tinsley mirror galvanometer G. The standard resistance of 5 ohms in series with the wire enabled the wire current to be evaluated, by measuring the voltage drop across the resistance with a Pye No. 7569 potentiometer P. The variations of R_w with i, obtained for the two separate lengths of wire, are shown in Fig. 3. From these curves the value of r_c was obtained, as indicated by Eqn (2).

The resistance–temperature calibration for the wires was obtained by immersing a 1½ in. length of wire, in an atmospheric environment, in a horizontal work tube in a Gallenkamp electric furnace. The wire resistance was found using a Wheatstone bridge and the temperature by measuring the e.m.f. from a sheathed platinum/platinum–13% rhodium thermocouple, also immersed in the furnace tube, with a Pye potentiometer.

The dependence of such a calibration on the chemical composition of the gaseous environment of the wire has been investigated by the authors (Bradley and Entwistle 1961). It was found that after immersion of platinum or platinum–10% rhodium wires in nitrogen, argon and air separately at temperatures above 1200° C and subsequent recalibration of the wires up to 1000° C, the temperature coefficients of resistance obtained depended upon the gas in which the wire was originally immersed at the higher temperature and upon the time of immersion. Also, after electrically heating a

plain wire *in vacuo* an increase in room temperature resistance was observed. Subsequent reheating in air caused the room temperature resistance to decrease to its original value. Quartz coated wires were not affected in the same way; they maintained a stable calibration. All these effects are probably due to gaseous adsorption or desorption by the metal.

For the work described it would have been better to calibrate the wires in a vacuum furnace, but this facility was not available. In order, therefore, to reduce to a minimum possible changes in the resistance–temperature relationship when heating *in vacuo*, the following procedure was adopted:

(i) The wire was heated *in vacuo* for the shortest possible time, just long enough to obtain the required readings.

(ii) The wire was frequently reheated electrically in air in order to re-adsorb any desorbed air, or oxygen.

(iii) After taking readings at the higher wire temperatures, where adsorption and desorption effects are more marked, points previously obtained at the lower temperatures were checked.

4. Emissivity of the wires

The curves shown in Fig. 3 enabled the resistance r_c to be obtained for different values of the current i, using Eqn (2). Knowing the wire resistance–temperature relationship, T_c could be deduced and the first term on the right-hand side

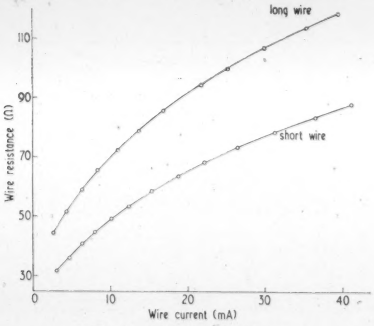


Fig. 3. Wire heating curves.

of Eqn (3) evaluated. If the radiant energy received by the wire is neglected then the second term on the right-hand side of Eqn (3) disappears. The value of the emissivity obtained on this assumption is called the apparent emissivity and the values of this, derived from the experimental readings, are listed for different wire temperatures in the Table. The determination of the true emissivity requires a knowledge of temperature $(T_b T_c)^{\frac{1}{2}}$. Using this expression for α, by a process of iterration, values of the true emissivity ϵ were

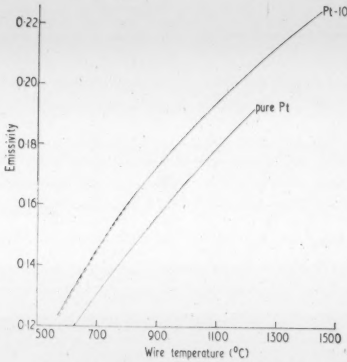


Fig. 4. Emissivity at different temperatures.

derived from the values of apparent emissivity. These values are listed in the Table and are shown by the full line curve in Fig. 4. At wire temperatures above 900° c the term containing α in Eqn (3) is small enough to make the two curves coincide. Also shown, for comparison, in Fig. 4 are the experimental values of the emissivity of pure platinum wire, obtained by Davisson and Weeks (1924). These workers compared their values of observed emissivities with theoretical values obtained from an expression deduced from Fresnel's laws of reflection and Maxwell's theory of electromagnetism. The expression

$$\epsilon = 0{\cdot}751(T\rho)^{1/2} - 0{\cdot}632(T\rho) + 0{\cdot}670(T\rho)^{3/2} - 0{\cdot}607\ (T\rho)^2 \quad (4)$$

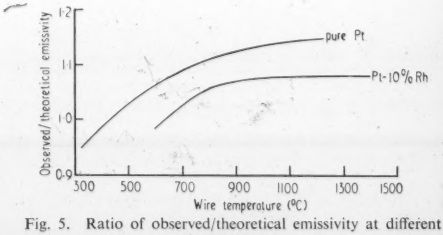


Fig. 5. Ratio of observed/theoretical emissivity at different temperatures.

Apparent and actual emissivities at different wire temperatures (°C)

Temp (°C)	600	700	800	900	1000	1100	1200	1300	1400	1450
Apparent ϵ	0·1276	0·1444	0·1597	0·1723	0·1837	0·1937	0·2032	0·2122	0·2206	0·2243
Actual ϵ	0·1287	0·1450	0·1599	0·1723	0·1837	0·1937	0·2032	0·2122	0·2206	0·2243

the value of absorptivity α at the appropriate wire temperature.

Figure 4 shows emissivities at different wire temperatures. The broken curve, based upon $\alpha = 0$, shows the variation of apparent emissivity. For a metal surface at absolute temperature T_c and for incident black body radiation from absolute temperature T_b, Eckert and Drake (1959) have shown, from electromagnetic theory, that the value of α for the surface is equal to the value of ϵ for the surface at the where ρ = wire resistivity in ohm cm, gives the emissivity as a function of temperature and resistivity.

Davisson and Weeks evaluated the ratio of observed/theoretical emissivities at the different wire temperatures and the relationship is shown in Fig. 5. The deviation of the ratio from unity was discussed by the authors. They suggested that, as the wire temperature increases, the shorter wavelengths make a relatively greater contribution to the energy

radiated and the resistivity for these frequencies becomes appreciably greater than for the zero frequency, assumed in the theoretical expression. At the lower temperatures the resistivity is less than that for zero frequency, possibly due to the response of the system to frequencies close to its natural frequency.

More recent theory shows quantitatively how, for a conductor, the spectral emissivity at the shorter wavelengths increases as a consequence of the limited response of the free electrons to the higher frequency field (Harrison 1960).

Using expression (4) for theoretical emissivity the ratio of observed/theoretical emissivity was evaluated for the platinum–10% rhodium wires used in the present work and these values are also shown in Fig. 5. The similarity of the two curves is apparent and suggests a method of estimating emissivities of platinum–rhodium alloys from their resistivities.

5. Experimental accuracy

The experimental method employed depends upon the central temperature of the long and short lengths of wire being the same, with the same electric current flowing. It would, however, be possible to have a length of wire so short that the central temperature was lowered by conduction to the main leads. A check was made to see if the assumption of identical central temperatures was justifiable, by calculating the temperature distribution along the wires from Eqn (1).

The equation was put into finite difference form and solved for different values of the current and a main lead temperature of 300° K using the relaxation method. In deriving the temperature distribution the experimental values of the wire resistance–temperature calibration were used. Furthermore, Eqn (3), used in the experimental determination of ϵ, can be rewritten

$$\frac{4\sigma}{d}(\epsilon T^4 - \alpha T_b^4) = \frac{4i^2 r}{J\pi d^2}$$

for a central temperature T and thus the third term in Eqn (1) may be found from the experimental readings as a function of T. In this way the experimental results were used to obtain the temperature distribution in order to check the validity of the original assumption of identical central temperatures. Values of the thermal conductivity of platinum–10% rhodium were not available through the required temperature range but the room temperature values of Barrat (1914) were extrapolated to higher temperatures, being guided by the temperature variation of the conductivity of platinum with temperature (Holm and Stormer 1930).

Two temperature distributions obtained in this way for wires with central temperatures of 1445° C and 647° C are shown in Fig. 1, the centre of the shorter wire used in the experiments being indicated by the line AA. The relatively greater end cooling effect at the lower central temperature is apparent. The temperature distribution with a central temperature of 647° C showed that, for the same current, the central temperature of the shorter wire was within 0·5° C of the central temperature of the longer wire. As the central wire temperature falls below 500° C this temperature difference increases appreciably and, for the wire lengths used in these experiments, this temperature is the lowest at which emissivities may be found.

The accuracy of the temperature measurements made is ultimately dependent upon the accuracy of the thermocouple reading in the calibration furnace. The thermocouple was calibrated against a platinum resistance thermometer up to 1000° C and was accurate to ±1 deg C. However, at the maximum temperature of 1450° C there was a possible error of ±0·25% of the reading. Consideration of this error, together with the possible errors in wire measurement, suggests that the maximum possible error in the value of the emissivity occurs at the highest temperature and is of the order of ±1·5%.

References

BARRAT, T., 1914, *Proc. Phys. Soc.*, **26**, 347.

BRADLEY, D., and ENTWISTLE, A. G., 1961, *Fourth Symposium on Temperature, Its Measurement and Control in Science and Industry* (New York: Reinhold), to be published.

DAVISSON, C., and WEEKS, J. R., JNR, 1924, *J. Opt. Soc. Amer.*, **8**, 581.

ECKERT, E. R. C., and DRAKE, R. M., 1959, *Heat and Mass Transfer* (New York, London: McGraw-Hill), p. 375.

HARRISON, T. R., 1960, *Radiation Pyrometry and its underlying Principles of Radiant Heat Transfer* (New York, London: John Wiley).

HOLM, R., and STORMER, R., 1930, *Siemens-Konzern*, **9**, 300, *International Critical Tables*, 1927, Vol. 2 (New York, London: McGraw-Hill), pp. 461, 467.

Observation accuracy in relation to the British Standards Institution's draft recommendation for industrial instrument scale design

by J. SPENCER, M.A., Department of Psychology, University of Bristol

MS. received 1st *September* 1961

Abstract

The correct specification of industrial instrument observation error in terms of magnitude and frequency of error is discussed and experimental results are presented for two examples of scale design from the British Standards Institution draft recommendations. These results show that the design performance (95% *of all observations shall have errors not exceeding* $\pm 1\%$ *of maximum scale value for a scale commencing at zero) is achieved providing that observers have a minimum of practice and are allowed a maximum of two seconds observation time under adequate illumination.*

1. Introduction

In a recent article in this *Journal*, Maddock (1961) has described the evidence and the conclusions which form the basis of the British Standards Institution (1961) draft recommendations on instrument design.

Whilst the evidence is impressive in volume, it presents a rather fragmentary picture of scale design and leaves considerable latitude for interpretation. This state of affairs is in part due to the fact that several design features of a good scale are not critical for normal industrial purposes. But a good part of this latitude arises because experimental investigations have not covered every possible combination of the variables involved in a given scale. This is for very obvious reasons and no one would suggest that, because of this, it is not possible to assemble a body of reasonable generalizations about desirable design features. But this does mean that such generalizations should be subjected to examination to see whether they are valid or in need of modification. It is the purpose of this note to describe an investigation of two designs based on the *Recommendations* to see whether they yield the observation accuracy predicted for them. It will be useful to precede this description by a brief discussion of what we mean by observation accuracy, so that the meaning of the results will be more readily appreciated.

2. Observation accuracy

When the physical accuracy of an instrument movement is specified, it is usually sufficient to state the range of the error in position of the index or pointer driven by the movement. For example, this is commonly expressed as $\pm x\%$ of full scale deflection. This value is a constant, to all intents and purposes, and very precise limiting predictions can be made about the true value represented by any given pointer displacement on the assumptions that no drift occurs and that no departure from design operating conditions is allowed to occur.

In the same way, the observation accuracy achievable with a given scale may be expressed as $\pm x\%$ of the maximum scale value (abbreviated to m.s.v.* in what follows) providing that a fundamentally necessary qualification is added which specifies what proportion of any given series of observations will, or are predicted to, lie within the $\pm x\%$ of m.s.v.

The qualification is necessitated by the characteristics of the observer who, by virtue of the complexity of his sensory and cerebral mechanisms, is never completely error free. Every so often, even when specially motivated to avoid it, he will misreport an indication. Furthermore, it is difficult to predict the magnitude of these errors. In general, errors may be conveniently divided into two classes, one class consisting of small errors and the other of large errors. The small errors arise through lack of sufficient 'resolving power' to achieve perfect visual discrimination of the exact position of the pointer in relation to scale marks. A second source of small errors is failure correctly to establish, or report, the small scale division occupied by the pointer. This involves partly visual and partly cerebral mechanisms. Large errors, which are much less frequent in practice, seem to occur as a result of a cerebral aberration and are little understood. For example, a typical example would be reporting an indication of 425 as 225. Good design can dramatically reduce the incidence of both large and small errors, although the latter always remain as long as the designer has to achieve compactness of scale as one of his objectives. Due to these various sources of error then, a typical observer during a series of observations on a given scale will closely approximate to a normal frequency distribution of error magnitude, with a small proportion of extreme error values. The mean of this error distribution is zero, or very close to it, with a standard deviation which varies according to the scale design, conditions of observation and the degree of practice of observers. The scale designer's problem may thus be considered as one of reducing the value of the standard deviation of observation error to a value approximately one-quarter the allowed error tolerance measured in scale units.

It is clear, therefore, that observation accuracy involves not only the physical magnitude of error in terms of scale

* m.s.v. is used in this article as an abbreviation for maximum scale value. It is assumed that the scale starts from zero. Maximum scale value is defined in B.S. 2643, Glossary of terms relating to the performance of measuring instruments, as follows: "the greatest value of the measured quantity which the scale is graduated to indicate".

units, but it also involves the frequency with which observations must lie within the allowed error tolerance. If the user requirement is that 90% of all observations taken must lie within ±2% of m.s.v. then a scale designed around an objective of 80% of observations lying within ±2% of m.s.v. will not be adequate. A simple analogy of this situation is perhaps that of selecting an electrical resistance for a circuit. It is necessary, but not sufficient, to know what ohm value is required. It is at least as important to know what watt rating is required, since the same ohm value is available in a large number of different watt ratings. For any given circuit application only one value of ohms and one value of watts will be optimal. The British Standards Institution Draft Recommendations suggest that a reasonable design objective is that 95% of any sample of observations by practised observers under specified observation conditions shall have errors not exceeding ±1% of m.s.v. for scales starting from zero. If the scale commences at a value other than zero, the error limits will be ±1% of scale range.

3. Experimental method

An apparatus was constructed which allowed an observer to see a dummy indicator for a period of two seconds when he pressed a switch which controlled a solenoid-operated shutter placed in front of the dummy indicator. The dummy indicator had a movable pointer which could be accurately set to known positions by the experimenter. After each exposure of the indicator, the experimenter reset the pointer in readiness for the next exposure. A series of ten exposures was made for each of the two scales used in this examination. Before each experimental series of exposures, the scale was shown to the observer and its features were described. He was told to ask questions about the scale if he was not sure about the markings and so on. After this, and before commencing the experimental series of observations, he was allowed to make five practice observations. During the experimental series, the observer reported as accurately as possible each indicator setting as soon as he could during or after its exposure.

The phrase 'as accurately as possible' was used deliberately in describing the task to the observer, to convey the idea that he was not limited by any arbitrary standard of approximation. He was told that reporting to the nearest scale unit (i.e. interpolation of a scale division into twentieths) was desirable, but that if he wished he could report to the nearest half of a scale unit (i.e. interpolation of a scale division into fortieths). It was further pointed out that he would probably regard the latter as guess-work, but that nevertheless he was to go ahead and guess. This type of advice was given because it is well known in psychological literature that very often what prevents a person achieving the high degree of accuracy he is capable of achieving is simply lack of confidence in reaching a decision. Instructions that guessing is allowed thus act as a restorative to confidence and allow best performances to be achieved. 'Second thoughts' were allowed, but only one exposure at a given indication was permitted.

This procedure was followed for a total of 24 observers. The order of presentation of designs was alternated and the settings selected to cover most of the scale range evenly and to provide even representation of unit digits in the indications.

4. Indicator scales

The two designs studied are illustrated in Fig. 1. The 0–600 scale was chosen because it is one of the least satisfactory of the compromises adopted in the British Standard

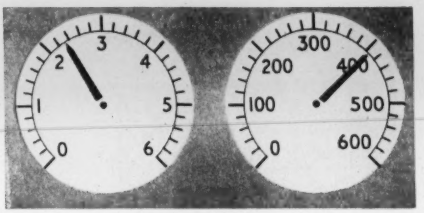


Fig. 1. Illustration of the scales investigated.

Institutions draft recommendations. It may therefore be regarded as representing the worst case for an examination. If it yields satisfactory results one can assume that the other recommended scale ranges would perform equally well, if not better.

The choice of the 0–6 scale was made because it was thought that accurate reading involving two places of decimals might lead to confusion and increased errors by observers. In both designs the scale base diameter was 2·10 in.

Four different designs of pointer were employed during the investigation but subsequent analysis showed that no differences in observation accuracy occurred due to this variable. For the purpose of this paper, therefore, this feature of the investigation will not be discussed.

5. Observation conditions

Observers sat comfortably in an office chair placed so that their eyes were 6 ft from and about 6 in. above the centre of the indicator.

The scale markings and digits were black on a white dial. The panel supporting the indicators was painted matt finish medium grey and was 25 in. wide by 16 in. high. Illumination was provided by two 40 w pearl bulbs in Anglepoise lampholders positioned on either side and 2 ft from the indicator; the luminance of white dial surface was 17·8 ft lamberts and that of the panel was 5·6 ft lamberts.

6. Observers

Twenty-four process operators from an oil refinery volunteered as observers. Their ages ranged from 21 years to 50 years, with a median age of 34 years.

They were not examined for visual capacity since the scale designs allow for variations in this respect.

7. Results

The overall results are illustrated by the graph, Fig. 2. This shows that between 88% and 90% of the observation sample was within the specified tolerance of ±1% of m.s.v. (i.e. within ±6 scale units or ±0·06 units of the true indication respectively). This is a little below the 95% specified by the B.S.I. draft recommendations. However, these observers, although familiar with instrument observations, were not practised in these designs. It was not possible, of course, to keep these men from their work for more than a minimum of time, so that a full-scale experiment designed to show improvement with practice could not be carried out. A check on the effect of practice was carried out, however, on one observer.

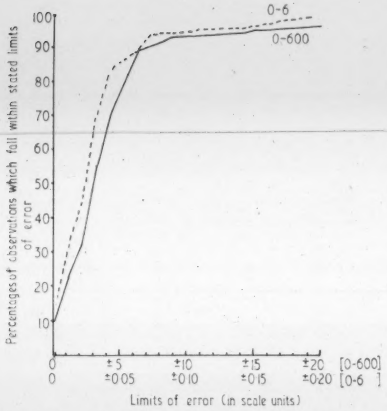


Fig. 2. Cumulative frequency distributions of observation errors for two scales (illustrated in Fig. 1). The graph shows, for example, that for the 0–600 scale, 89% of observations are at least as accurate as ±6 scale units.

This observer was selected as being of average performance from the results obtained in the experiment. Permission was obtained to use him as an observer for a second period. During this second series, he made a further 80 observations on the 0–600 indicator. The graph, Fig. 3, shows the change

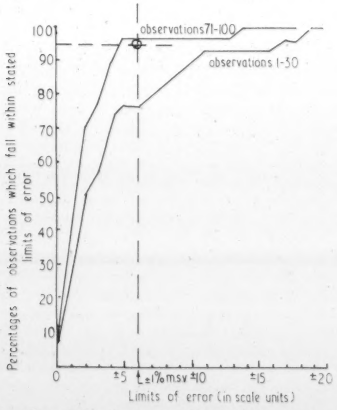


Fig. 3. Cumulative frequency distributions of observation errors at two stages of practice for one observer.

in performance between the first 30 and the last 30 observations. It can be seen that after 70 observations he is achieving a 95% proportion of correct readings. Consequently, it may be said that for purposes of definition, a practised observer is one who has made a total of at least 100 observations on a particular scale design. The difference in r.m.s. error between the first and last batches of 30 observations is very significant statistically, so that the claim can be made that these results are evidence of genuine improvement and not just a fortuitous difference.

The description of results can be completed by a brief survey of the gross errors that occurred in the main experiment. Out of a total of 480 observations (2 scales × 10 observations × 24 observers) 24 (i.e. 5%) were in error by more than ±1½% of m.s.v. Of these 24, there were four errors greater than ±10% of m.s.v. (i.e. ±60 or ±0·6 scale units) and all of them involved mislocating the numbered segment of the scale, so that they were errors of ±100 or ±1·0 scale units respectively. It is significant that two of the four occurred when the pointer obscured part of the nearest scale number group. These four errors then correspond to those described earlier as due to a 'cerebral aberration' and constitute a little under 1% of all observations.

The remaining 20 errors were small and 16 of them were errors of almost exactly ±10 or ±0·1 scale units. The failure here is due to mistaking the small scale divisions as equal to 10 instead of 20 scale units. This is a common source of error which is difficult to combat, because if minor divisions are shown for every ten scale units, the observer is then much more likely to 'mis-count' the number of small scale divisions. Hence the cure exacts a penalty of errors of a different cause, and in most cases it will be found that these latter will be more frequent than the original errors.

8. Conclusions

It is concluded that these two scale designs yield the specified performance claimed for them in the British Standards Institution draft recommendations. That is, 95% of all observations taken will be within ±1% of m.s.v. of the true indicated value. This performance is obtained in the following conditions: that a luminosity of at least 5 ft lamberts on white surfaces prevails, that a maximum of 2 seconds is available for observation and that the observers are practised on the scale (the evidence presented here suggests that about 100 observations will be sufficient, but minimum, amount of practice) and trying to observe it as accurately as possible.

Acknowledgments

I would like to express my gratitude to Mr. G. W. Amphlett of Messrs. Payne and Griffiths Ltd., and to Mr. A. Throp of Messrs. Matthew Prior and Son Ltd., who jointly provided the components of the dummy indicators.

Also I would thank all those members of the British Petroleum Refinery, Llandarcy, who either assisted in making the arrangements or volunteered to act as observers, and without whose generous help the experiment could not have been carried out.

Permission to publish this paper has been kindly given by the Director of the British Standards Institution and the Department of Scientific and Industrial Research.

References

British Standards Institution, 1961. *Draft British Standard Recommendations for the Graduation of Industrial Instruments for Quantitive Measurements* AB(INE) 1661, 19th May, 1961 (London: British Standards Institution).

MADDOCK, A. J., 1961, *Brit. J. Appl. Phys.*, **12**, 33.

Dielectric behaviour of mercuric oxide

The variations of the permittivity and of the loss tangent of the red and yellow modifications of high purity HgO with temperature, in the interval between −78·5 and +120° c were determined at frequencies between 50 c/s to 5 Mc/s, using compressed powder specimens (grain size less than 20 μ) with allowance for porosity (Bottcher 1942, 1945). It has been found that the apparent dielectric constant of the red oxide shows larger dispersion than the yellow oxide, but the two oxides converge to a value in the vicinity of 9 at low temperatures and high frequencies (Fig. 1). The a.c. conductivity (Fig. 2) indicates that the red oxide may contain, possibly more than the yellow, impurity centres which become ionized below room temperature, and its normally higher dielectric constant is probably associated with this ionization. The loss mechanism appears to be a Maxwell–Wagner type of electronic character, and is possibly due to stoichiometric excess of mercury, a conclusion which is supported by the sign of the thermoelectric force which suggests that, like ZnO (for example, Hahn 1951), HgO is an n-type semiconductor.

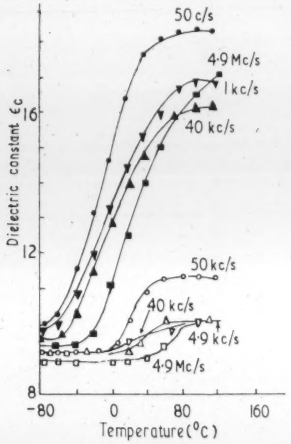


Fig. 1. Temperature dependence of the apparent dielectric constant of HgO, with the frequency as a parameter. Solid symbols indicate red HgO, open symbols yellow HgO. The true permittivity at static frequencies is probably near 9.

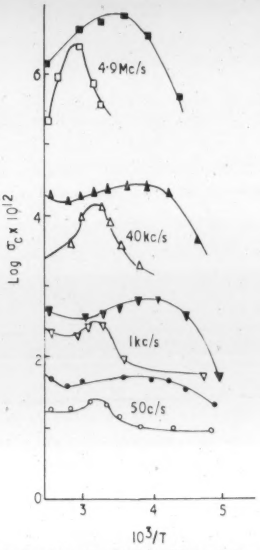


Fig. 2. Relation between the logarithm of the a.c. conductivity and the reciprocal absolute temperature. The units of σ_c are ohm^{-1} cm^{-1}.

This work is a part of a programme planned by Professor A. R. Tourky, Head of the National Research Centre of Cairo, United Arab Republic, for investigating the physico-chemical properties of some metallic oxides.

Physics Department, University of Cairo, United Arab Republic. — Y. L. Yousef, S. A. Zaher

National Research Centre, Cairo, United Arab Republic. — Z. Hanafy, T. M. Salem

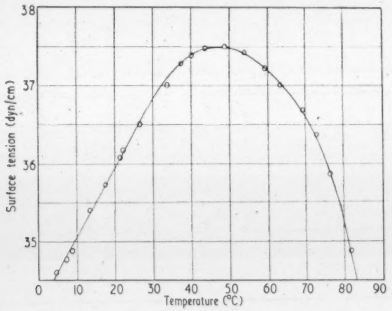
19*th September* 1961

Bottcher, C. J., 1945, *Rec. Trav. Chim.*, **64**, 47.
—— 1942, *Physica*, **9**, 937.
Hahn, E. E., 1951, *J. Appl. Phys.*, **22**, 855.

Anomalous variation of surface tension

In the course of surface tension measurements on a proprietary fluid* the authors have found an anomalous variation of surface tension with temperature, shown in the figure.

The fluid consisted of a glycol–water mixture in approximately equal proportions, the glycol being made up of several lower members in addition to a high polymer glycol.



Variation with temperature of surface tension of the fluid.

The surface tension was measured with a du Noüy ring tensiometer to an accuracy of $\pm 0{\cdot}1$ dyn/cm over the temperature range 5° C–80° C. Measurements were taken on several samples of the fluid and whilst the specific values recorded differed slightly from sample to sample, the shape of the curve was consistent in each case with the maximum appearing at about 45° C.

A survey of the literature has revealed that a reversal of temperature coefficient of surface tension with change of temperature has been observed in the cases of liquid crystals (Adam 1946, Ferguson 1943). All other cases noted, e.g. polyethylene glycols and their derivatives (Gallaugher and Hibbert 1937) and alcohol–water mixtures of varying proportions (Bonnell, Byman and Keyes 1940) as well as many other liquids show consistently a negative temperature coefficient. This has been found also by the authors for ethylene glycol–water mixtures containing from 100% to 50% glycol. Altenburg (1952) neither found nor expected anomalous variations of surface tension with temperature for ethylene glycols.

* The fluid used is commercially known as Houghtosafe 271 and the authors acknowledge gratefully the co-operation of the suppliers, Messrs. Edgar Vaughan and Co. Ltd., Birmingham.

Two possible factors which may suggest lines of investigation are now considered.

Firstly this anomalous rise of surface tension with temperature may be due to a variation of association of the polymer glycols with rise of temperature. Presumably this cannot be due entirely to the lower member glycols in view of the authors' work on ethylene glycol–water mixtures. Unfortunately evidence shows that the relation of degree of association to surface tension, by means of the Eötvös–Ramsay Shields equation, for example, is open to question (Adam 1946). Further Mitra and Sanyal (1958) have derived on theoretical grounds the equation $\log S = A - B/(T_c - T)$ where S is the surface tension, T is the temperature, T_c the critical temperature, and A and B are constants. The equation was shown to hold in cases of both associated and non-associated liquids and to describe a falling surface tension with increase of temperature. It is clear that a law of this form could not describe both a rise and fall of surface tension.

Secondly, since the positive temperature coefficients observed in the case of the liquid crystals occur during the change from anisotropic to isotropic states, it may be that some orientation in the cooler mixture is destroyed by a rise in temperature. This line of thought has some support in the paper of Gallaugher and Hibbert (1937).

It may be noteworthy that Skala (1912) reported a reversal of temperature coefficient for glycerine–water mixtures but this change was due to variation of mixture proportion and not temperature. The present authors did not at first rule out selective evaporation, for example, as a possible link here but it was found that cooling of the fluid used resulted in a reversal of the surface tension trend and so evaporation could not be causing the effect noted.

Physics Department,
College of Advanced Technology,
Birmingham.

R. W. Cotterhill
R. K. Fitch

4th October 1961

Adam, N. K., 1946, *The Physics and Chemistry of Surfaces* (Oxford University Press), p. 157.

Altenburg, K., 1952, *Z. Phys. Chem.*, (*Lpz.*), **201**, 91.

Bonnell, W. S., Byman, L., and Keyes, D. B., 1940, *Ind. Engg Chem.*, **32**, 532.

Ferguson, A., 1943, *Endeavour*, **2**, 24.

Gallaugher, A. F., and Hibbert, H., 1937, *Amer. Chem. Soc. J.*, **59**, 2514.

Mitra, S. S., and Sanyal, N. K., 1958, *Z. Phys. Chem.*, (*Lpz.*), **208**, 257.

Skala, J., 1912, *Akad. Wiss. Ber.*, **121** (2a), 1213.

Thermal expansion of bismuth telluride

In a recent examination of the properties of bismuth telluride at high temperatures, it became desirable to know the thermal expansion coefficient as a function of temperature in both the *a* and *c* directions in the crystal.

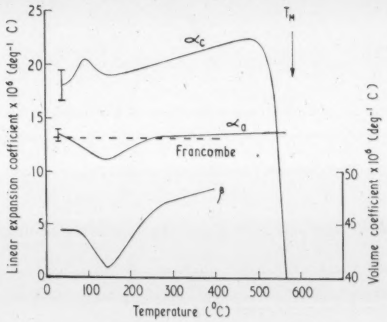
Specimens were prepared by fusion of stoichiometric proportions of the elements in vacuum and subsequent zone melting to align the crystallites in one direction. It was then possible to obtain single crystal specimens from this ingot, which were used to determine the expansion coefficients. Parallel to the cleavage plans these were about 3 cm in length and perpendicular to this direction 0·9 cm.

The linear coefficient of expansion α was measured relative to quartz, using a conventional push-rod technique with an optical lever capable of detecting changes in specimen length of less than 10^{-4} cm. In this way the results shown in the Figure were obtained.

These results differ markedly from those of Francombe (1958). However, his values are based on a limited number of points in this temperature range obtained by an x-ray technique. If a mean value of the expansion coefficient is found from the unit cell data of his work, its magnitude is in much better agreement with these results. These mean values are also shown in the Figure.

The sudden decrease in α_c near the melting point was reproducible, although the specimen length changed slightly each time this temperature range was covered. It is thought that slip may be taking place between the planes and a certain amount of distortion was found on the sides of the specimen on removal from the apparatus. This pre-melting phenomenon may be associated with the anomalously high value of the heat capacity in the same temperature range which have been observed by Bolling (to be published). In addition, further support has been found by Hodgson (private communication, 1961) using a differential thermal analysis method.

The linear expansion coefficient observations have been used to derive values for the volume coefficient of expansion and these are also shown in the Figure. The low temperature minimum in this parameter has been confirmed by Packwood (private communication, 1961) using a volume dilatometer.



Expansion coefficient of bismuth telluride.

I wish to thank both R. Hodgson and R. Packwood of Birmingham University for making the measurements referred to under their names in the above article and also Professor J. Sayers of the Electron Physics Department of that University, where the work was carried out.

Department of Physics,
Science Laboratories,
South Road,
Durham City.

K. N. R. TAYLOR

5th October 1961, *in revised form* 31*st October* 1961

FRANCOMBE, M. H., 1958, *Brit. J. Appl. Phys.*, **9**, 415.

New books

Fortschritte der Hochfrequenztechnik, Band 5. By M. Strutt and F. Vilbig. Translated by F. Rühmann. (Frankfurt: Akademische Verlagsgesellschaft, 1960.) Pp. x + 412.

This is the fifth volume in a series entitled 'Progress in High Frequency Techniques', the preceding volumes having appeared in 1941, 1943, 1954 and 1959. The later volumes consist of about eight chapters on different aspects of high frequency technology, the qualification 'high frequency' being taken in its broad sense of radio frequency. The chapters are by different authors, each an authority on the subject matter concerned. This is an advantage that has been increased in the last two volumes by widening the source of authors to include other nationalities. In volume five, three chapters are written in German and five in English, but abstracts and figure legends are given in both languages.

Two separate chapters of a related nature are of interest to those concerned with the transmission of speech and television signals over low capacity channels. Emphasis is primarily on coding techniques for band compression, although the subjective aspects are not ignored. Theoretical discussions indicate the degree of bandwidth compression that should be achievable and these are followed by descriptions of some methods that have been tried. Attention is drawn to the essential complication of the apparatus and to the reduced entertainment value. Communication systems are represented by a chapter on recent advances in digital communication, covering the theory of signal error probabilities due to additive Gaussian noise and to Rayleigh fading. There is discussion of diversity methods and error-correcting codes but little of circuit techniques. Three chapters lie in the field of electron tube design. One shows how to determine approximately electron trajectories in electric and magnetic fields taking account of space charge forces, a problem which becomes important as current densities increase in relation to tube potentials, a second describes the operation of charge controlled storage tubes, and the third deals with the design of millimetre wave valves along conventional lines. Until one or more of the new principles of millimetre-wave generation reported in the literature actually materializes into a practical device, there is a need for ingenuity in straining those techniques already in use at slightly longer wavelengths. The design of non-reflecting surfaces at centimetric wavelengths is treated quite thoroughly, proceeding from an exposition of basic principles to details of experimental results. The volume would hardly be representative without some reference to semiconductor progress and the particular aspect reviewed is the development of high frequency transistors of various types. The influence of the various parameters on the frequency limitation is illustrated by mathematical and physical descriptions. Tunnel diodes are also included.

In each subject progress is reviewed over the last few years. The development is presented logically and sufficient theory is given so that the reader can learn something of each as well as become *au fait* with the present state of the research. The decision to publish this book annually should be welcomed.

R. DALZIEL

The encyclopaedia of spectroscopy. Edited by G. L. Clark. (New York: Reinhold Publishing Corporation; London: Chapman and Hall, 1961.) Pp. xvi + 787. Price 200*s*.

This impressive book should not have been called an encyclopaedia. If it were one, the coverage would be complete and redundancies avoided. The articles would be concise and appropriate in length to the importance of the topic concerned. The matter would be logically arranged so that one could find a subject without difficulty.

Actually it is a collection of articles, some excellent, others poor. One article ('Differential thermal analysis') seems to have drifted in from another encyclopaedia. Omission of this, and stricter editing, might have left room for the longitudinal observation of the Zeeman Effect, with the consequent circular polarization, for going beyond L–S coupling and mentioning j–j coupling and others, for giving more than eighty words to the Stark Effect (and that in the article on the Zeeman Effect), for the theory of x-ray spectroscopy, and for many other subjects. Fourteen and a half pages of 'Paint Products' is long under the circumstances.

The modernity of the book is evidenced by an article on 'Optical Masers', said to have been derived from two news items, one in the *New Scientist* and one in the *Denver Post*.

The publisher states that "the articles are not merely abstracts or highly condensed versions of original manuscripts. Each article was especially written for inclusion in this volume and covers every major aspect of its topic." Further, the publisher claims that there were "over 160 internationally recognized contributors". In fact seventeen articles have footnotes acknowledging either previous publication or their condensation from articles previously published. The list of contributors in the book contains 121 names.

It is dear at £10; half the book at half the price, carefully selected and edited, would be good value. A. C. MENZIES

Progress in aeronautical sciences, Vol. 1. Edited by A. FERRI, D. KÜCHEMANN and L. H. G. STERNE. (London, New York, Paris, Oxford: Pergamon Press, 1961.) Pp. xix + 280. Price 80*s*.

This is the first of a series of volumes to be published annually, each of which will contain a number of review articles of aeronautical interest. In the words of the preface, it is hoped that these articles "will provide the specialist reader with an orderly but concise summary of recent work and so relieve him of the need for excessive recourse to original papers; and the general reader with an opportunity to become painlessly but usefully informed of recent developments in fields other than his own". The volumes are thus to a large degree an aeronautical counterpart of the 'Reports of Progress in Physics' or the 'Reviews of Modern Physics'.

It is likely that a book of this type will consist of several more or less unrelated articles and the present volume is no exception. In the first article, E. C. Maskell of the R.A.E., Farnborough, discusses the basic principles of aerodynamic design, pointing out that present-day aircraft shapes, with attached flow over nearly all surfaces, are not the only possible forms which give acceptable, steady flows. Though well written, this article is very short (7 pages) and serves to whet rather than satisfy the appetite for aerodynamic philosophy. The second article is by Robert Legendre and sets out, in French, a method of designing compressor and turbine blades, particularly when the flow about the blades is transonic. The third article considers a related problem, that of the use of potential flow methods for the design of wings for supersonic flight. It too is in French and is written by M. Fenain of O.N.E.R.A. In both these papers the subject is treated strictly from the theoretical standpoint; no comparison with experiment is made.

The fourth article is an excellent account in German by E. Becker of Göttingen of the recent work on unsteady boundary layers, particularly those developing in the short-duration flows that occur in shock tubes and hypersonic test facilities. Existing theories are reviewed and compared with experiment and there is an extensive bibliography. For physicists, the fifth article is perhaps the most interesting. By F. A. Goldsworthy of Manchester University, it discusses in some detail the dynamics of an ionized gas. The rapid progress of sustained and ballistic flight has forced aerodynamicists to regard air as a real, and not a perfect, gas and hence to become concerned with the details of dissociation and ionization processes. The flow of ionized gases is of course also of interest to those working in many other branches of physics, from kinetic gas theory to stellar explosions. Goldsworthy's review of this comparatively new subject is therefore very welcome, the more so since it sets out existing knowledge in a clear and concise manner.

The final two articles originate from the R.A.E. In one D. Williams considers the structural problems encountered in the design of aircraft pressure cabins. In the other, C. H. E. Warren and D. G. Randall set out the theory of the sonic bangs produced on the ground by aircraft flying at supersonic speeds. This article is useful in clearing up much of the mystery surrounding this topic and in showing that the observed effects may be explained quite rationally.

The overall impression is of a useful, if rather mixed collection of papers, of a sufficiently high standard to allow forthcoming volumes to be awaited with interest, though perhaps more by the aerodynamicist than the physicist. The printing and layout of the book are very good.

E. W. E. ROGERS

Semiconductor devices and applications. By R. A. GREINER. (London: McGraw-Hill, 1961.) Pp. xiv + 491. Price 97*s*.

Almost every author who writes a book on applications of semiconductors (and there are many) seems to feel that he must devote a fair portion of the book to a discussion of the basic physics of semiconductors. This is done apparently in a useless striving for completeness. It might not have been unreasonable when the physics of semiconductors was little known but is hardly justified now. Quite a number of both elementary and advanced accounts of semiconductor physics are now available and most of the potential readers of this volume are likely already to have a book on the subject; they will rightly feel that they do not need to pay again for the material in the first five chapters of the present book in order to have the rest; the price of the book is, in any case, rather high. Moreover, because of the need for compression, the author has been unable to present an entirely satisfactory account of the subject; indeed, in striving for conciseness and simplification he not infrequently misleads. The book could well have begun with Chapter 6 which is a very sketchy account of electrical contacts. Chapter 7 gives an account of p–n junctions which play a very important part in all semiconductor applications; the phenomenon of tunnelling is described but other important junctions like the $p–p^+$ and $n–n^+$ junctions are hardly mentioned.

The rest of the book is devoted mainly to the use of transistors and diodes in electronic circuits and very little is said about many other uses of semiconductors; to this extent the title is misleading. The next few chapters discuss the practical characteristics of transistors and diodes. There follows a discussion of transistor amplifiers for small signals and for higher powers, both a.c. and d.c. types being dealt with in some detail. Power supplies using junction rectifiers are then described, and stable oscillators using transistors. The last three chapters are devoted to switching circuits of the type used a great deal in electronic digital computers. The treatment follows a logical sequence suitable for a text book for engineering students who are more directly interested in the practical side of electronic circuits. The author rightly believes that an appreciation of the physics underlying the device operation is essential and has tried to present his material in such a way as to give some insight into the underlying physics.

A number of practical problems for solution are given at the end and these are chosen so as to consolidate the students' knowledge of the theory and also to make them aware of the practical and numerical aspects of the subject. The book is well produced in the well-known style of the publishers and is extensively illustrated. As stated above, it seems expensive for its contents, much of the material being available elsewhere. It does, however, form the basis of a logical course on transistor circuits. The rapid development of the practical forms of transistor and also of the techniques of using them means that it cannot long remain up to date, but at the moment it represents current practice fairly well.

R. A. SMITH

Flow of Fluids through Porous Materials. By R. E. COLLINS. (London: Chapman and Hall, 1961.) Pp. x + 270. 100*s*.

This useful volume reviews a field which has become of great interest to physicists and physical chemists, and which has long been of importance in petroleum technology, hydrology and soil science. The fundamental basis is potential theory, so that the subject is closely related to the flow of heat and diffusion of matter. This book is written more from the viewpoint of the petroleum technologist and hydrologist, and is perhaps at its best when considering simultaneous flows of miscible and of immiscible fluids, flow with change of phase and related problems.

In his preface, the author states that there are only two other books in the English language dealing with flow in porous materials, respectively by Muskat and by Scheidegger. However, there is also the monograph by Carman (*Flow of Gases through Porous Media* (London: Butterworths)) and the monograph-length article by Hubbert (*J. Geol.*, 1940, **48**, 785–944) which, in the reviewer's opinion, constitutes one of the best analyses of the scope of Darcy's law so far written. Indeed, although it is to the point, the bibliography is rather meagre, in view of the great wealth of relevant literature.

The new monograph comprises ten chapters, dealing with structure and properties of porous materials; statics of fluids in porous media; theory of flow; steady, homogeneous laminar flow; transient laminar flow; flow of immiscible fluids; moving boundary problems and deposition of solids; flow of miscible fluids; theory of models; and flow with change of phase.

A number of printer's errors appear, among which are the following: p. 31, 'sandpoint' for 'standpoint'; p. 32, in the ordinate of figure 2–6, K/ϕ should be $\sqrt{(K/\phi)}$; p. 49, equation 3.4, 'g' is not a subscript; p. 62, equation 3.45 should read '$p_\omega - p_g = p_{\omega g}$' not '$p_\omega = p_g = p_{\omega g}$'. However, these are obvious misprints which will not mislead. The book provides a useful addition to the literature of flow in porous materials.

R. M. BARRER

Notes and comments

Notes for the Preparation of Papers for the *British Journal of Applied Physics*

The following instructions refer to the main points which have to be considered in the preparation of a paper for publication, and authors offering papers to the *British Journal of Applied Physics* for publication are asked to conform to these recommendations.

Manuscripts

Manuscripts should be typed in *double* spacing on paper not wider than 8 in. and not longer than 13 in. Only one side of the paper should be used, and a margin of 1–1½ in. should be left. As alterations in the text cannot be allowed once the paper is set up in type, authors should aim at absolute clarity of meaning and of typing, and should check the typescript carefully before submission. All manuscripts should be submitted in duplicate, and in addition to the fair copies, a set of small copies or prints of the diagrams (not larger than foolscap) must be attached to each MS. (This enables MSS. to be sent to two referees at the same time and so assists rapid publication.)

Abstract. An abstract is printed at the beginning of the paper immediately after the title, name(s) of author(s), and place of employment of author(s). Two extra copies of the 'title and abstract' page are required for the records.

Mathematics. It is not necessary to give detailed derivations of mathematical expressions and formulae in a published paper when the work is straightforward; it is quite sufficient to indicate the method of treatment and the final results.

References. In the text bibliographical references are made by giving the name of the author and the year of publication in brackets, e.g. (Jones 1942), and details are given in the last section, 'References', where the references are arranged in alphabetical order of authors' names and in date order for each author.

Drawings

Drawings should be in Indian ink on tracing cloth, tracing paper or white card, with lettering in soft or blue pencil; lettering of a size suitable for reduction will be inserted by our draughtsman. The drawings should in general be sufficiently large to allow of reduction in printing, and the lines should therefore be bold; the frame lines of graphs should be slightly finer than those of the plotted curves.

Full details are available in the Institute and Society's 'Notes for Authors', obtainable from the Editor and Deputy Secretary, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London S.W.7.

Papers for publication in the *British Journal of Applied Physics* should be sent to the Editor and Deputy Secretary at the same address.

Centenary of Plastics 1862–1962

We have received notice that a joint committee of the Plastics Institute and the British Plastics Federation is to organize a number of events in connection with a Centenary of Plastics 1862–1962. These will include the publication of a book giving the story of plastics and containing a detailed history of celluloid, contributions in the national and technical press and on radio and television, open days in firms in the plastics industry and special meetings and lectures throughout the centenary year culminating in Interplas 1963, the International Plastics Exhibition to be held at Olympia from 12–22 June.

Further information may be obtained from the Plastics Institute, 6 Mandeville Road, London, W.1.

Journal of Scientific Instruments

Contents of the December issue

CONFERENCE REPORT

International Conference on Components and Materials used in Electronic Engineering—London, June 1961. By P. R. Bardell.

PAPERS

Distance measurement by means of a light ray modulated at a microwave frequency. By K. D. Froome and R. H. Bradsell.
Robust helical spring microbalance. By D. Griffiths.
Simple form of balance for the rapid measurement of magnetic susceptibilities. By E. V. Smith.
Metal cell for use in vapour pressure measurements by the effusion method. By S. Blairs, R. A. J. Shelton and R. Unsworth.
An improved field discriminator for magnetic field stabilization. By K. H. Marsden.
Self-contained integrating irradiance meter for use underwater. By L. Draper.
X-ray diffractometer for automatic operation. By W. A. Wooster and A. M. Wooster.
Low-velocity gas–liquid impinger for the continuous estimation of sulphur dioxide and other atmospheric pollutants. By T. Nash.
Direct-reading balances for percentage moisture content. By J. G. Downes, B. H. Mackay, L. G. Bellamy and V. D. Burgmann.
Measurement of rapid temperature changes by thermocouples. By J. A. Sirs.
Method for obtaining minimum noise in a general purpose hot wire amplifier. By G. N. V. Rao.
Safe universal mounting and alignment system for x-ray goniometers. By W. Hughes and C. A. Taylor.
Shock tube β-ray densitometer. By M. McChesney.
Construction and operation of a simple flame-ionization detector for gas chromatography. By D. J. Morgan.
Tripping a 1 MV single-stage impulse generator. By T. E. Broadbent and R. Cooper.

RESEARCH NOTES

Diamond knife trough for the ultramicrotome. By R. Phillips and J. H. Lucas.

LABORATORY AND WORKSHOP NOTES

Two megacycle ultrasonic interferometer. By J. H. Andreae and P. D. Edmonds.
Magnetically operated shutters for use in vacuum systems. By F. S. Feates.
Experimental microbarograph. By G. W. Gardiner.
Universal crystal mount for electron paramagnetic resonance measurements. By G. Roberts and W. Derbyshire.
Diode leak-detector control unit. By D. Allenden.
Stability of the e.m.f. of cells under continuous low-drain conditions. By J. Goddard.
Temperature rise in a barrier-layer photocell exposed to sunlight. By D. Paix.
Isotopic discrimination introduced by electron multiplier detectors in mass spectrometry. By B. L. Campbell and R. N. Whittem.
Simple device for digitalizing oscilloscope recordings on to punched tape. By I. M. French.

NOTES AND NEWS

New instruments, materials and tools Notes and comments

THIS JOURNAL is produced monthly by The Institute of Physics and The Physical Society, in London. It deals with all branches of applied physics (including theory and technique). All rights reserved. Responsibility for the statements contained herein attaches only to the writers.

EDITORIAL MATTER. Communications concerning editorial matter should be addressed to the Editor, The Institute of Physics and The Physical Society, 1 Lowther Gardens, Prince Consort Road, London, S.W.7. (*Telephone:* Kensington 0048.) Prospective authors are invited to prepare their scripts in accordance with the *Notes for Authors*.

REPRODUCTION. The Institute of Physics and The Physical Society is a signatory to The Royal Society's Fair Copying Declaration. Details may be obtained upon application from The Royal Society, London, W.1.

ADVERTISEMENTS. Communications concerning advertisements should be addressed to the agents, Messrs. George Jackson (Fleet St.) Ltd., Cliffords Inn, Fleet Street, London, E.C.4. (*Telephone:* Holborn 3611–2.)

SUBSCRIPTION RATES. A new volume commences each January. The charge is £6 per volume ($17 U.S.A.), including index (post paid), payable in advance. Single parts, so far as available, may be purchased at 12*s.* 6*d.* each ($1.75 U.S.A.), post paid, cash with order. Orders should be sent to The Institute of Physics and The Physical Society, 47 Belgrave Square, London, S.W.1, or to any bookseller.

CLAIMS FOR MISSING JOURNALS. Claims from regular subscribers to this *Journal* for missing numbers will only be considered if received within 60 days of the date of mailing plus normal outward time of transit and time for lodging the claim. Losses attributable to failure to notify a change of address or to similar omissions will not be considered.

PRINTED IN GREAT BRITAIN BY UNWIN BROTHERS LIMITED, WOKING AND LONDON